RECOGNITION MEMORY WITH AND WITHOUT RETRIEVAL OF CONTEXT: STUDIES WITH EVENT-RELATED POTENTIALS

Edward L. Wilding

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1996

Recognition memory with and without retrieval of context: Studies with event-related potentials

E.L. Wilding
December 1995

A thesis submitted to the University of St. Andrews for the degree of Doctor of Philosophy in the School of Psychology

Declaration for the Degree of Ph.D.

(i) I Edward Wilding hereby certify that this thesis, which is approximately 60,000 words in length, has been written by me, that it is a record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 18/12/95 signature of candidate

(ii) I was admitted as a research student in January 1993 and as a candidate for the degree of Doctor of Philosophy in January 1993; the higher study for which this is a record was carried out in the University of St. Andrews between 1993 and 1995.

date 18/12/95 signature of candidate

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date signature of supervisor

Thesis Copyright Declaration

Acknowledgements

I would like to thank Mick Rugg, my supervisor, for his excellent and balanced support over the past three years. I would like to thank Mike Doyle for the technical and practical assistance that he has given me, and the suggestions he made which contributed to the direction that the work in this thesis has taken.

I would like to thank Astrid Schloerscheidt and Kevin Allan for their comments on previous versions of this thesis, and Kevin in particular for many fruitful, if initially confusing, discussions.

Finally, I would like to acknowledge the continuing support of family and friends, and in particular Louise, for her patience and her help over the past three years.

Abstract

In six experiments event-related potentials (ERPs) were recorded while subjects performed modified recognition memory tests. All experiments consisted of an initial study phase in which subjects studied words which were presented in one of two contexts. In a subsequent test phase subjects discriminated between old and new items, and between old items which had been presented in one of the two contexts at study. The design of these experiments permitted a comparison of three critical classes of ERPs: those to words correctly judged new, and those to words correctly judged old which were either correctly or incorrectly assigned to study context.

All six experiments revealed reliable differences between the ERPs to correctly identified old and new words. In experiments 3-6 the analyses of the differences between the ERPs to correctly identified old and new words revealed two topographically and temporally dissociable modulations. The first of these was maximal at parietal sites over the 500-900 msec time window, and was larger over the left hemisphere than over the right. The second modulation was more temporally extended, maximal at frontal scalp locations, and displayed a right-greater than-left hemisphere asymmetry. Both of these modulations were of greater magnitude for words which were correctly assigned to study context. These findings are consistent with the view that multiple neural systems contribute to memory for context. The experimental findings are discussed in relation to theories of the relationship between memory for prior occurrences, and memory for contextual details of those occurrences.

Table of Contents

Chapter 1

1 Recognition with and without retrieval of study context

1.1 *Forms of memory*

It is uncontroversial to state that human memory is not unitary. Perhaps the most widely recognised distinction is between short-term (working) memory, and long-term memory (Atkinson and Shiffrin, 1968; Baddeley and Hitch, 1974; Milner, 1966; Shallice and Warrington, 1970; Waugh and Norman, 1965; Wickelgren, 1968). The former refers to a limited capacity system which stores material for relatively short periods of time. The latter has in principle no limits on storage capacity, and material may be stored indefinitely.

Within the long-term memory system a number of further distinctions have been identified. The following sections briefly review some of these distinctions, borrowing principally, but not exclusively, from the classification schemes due to Tulving (1993a), and Squire (Squire, Knowlton and Musen, 1993; Squire and Zola-Morgan, 1991). These frameworks have been selected primarily for descriptive purposes, and do not reflect a preferred theoretical orientation for the work contained in this thesis.

1.11 *Implicit and explicit memory*

One distinction between forms of long-term memory which is the focus of much contemporary research is between *implicit* and *explicit* memory (for definitions and important precursors see Cohen and Squire, 1980; Graf and Schacter, 1985; Graf, Squire and Mandler, 1984; Jacoby and Witherspoon, 1982; Tulving, 1983). The former refers to those forms of memory where prior experiences can affect behaviour without awareness of the influence of those experiences. By contrast, explicit memory refers to retrieval of information that is accompanied by awareness of the products of retrieval.

It has been proposed that implicit memory incorporates a number of different forms of memory, including the acquisition of motor skills and habits, simple classical conditioning, some forms of associative learning, and various forms of *priming* phenomena (Squire *et al.*, 1993; Squire and Zola-Morgan, 1991; Tulving, 1993a). This latter form of implicit memory refers to facilitations in task performance due to the influence of a prior learning episode. For instance, prior exposure to words facilitates reaction time in a subsequent lexical decision task (Scarborough, Gerard and Cortese, 1979). It is this form of implicit memory which is of most relevance to the theoretical issues which are addressed in this thesis.

Squire and colleagues (Squire, 1992; Squire and Zola-Morgan, 1991) divide explicit memory into *semantic* and *episodic* components, a distinction first introduced by Tulving (1972, 1983). The former refers to general and factual knowledge, whereas the latter, also referred to as autobiographical memory, refers to memory for specific past events and occurrences. Whilst a pragmatic distinction between semantic and episodic memory is generally accepted, there is considerable disagreement over the functional and neuroanatomical relationship between these forms of memory (Kinsbourne and Wood, 1975; McKoon, Ratcliff and Dell, 1986; Ratcliff and McKoon, 1986; Squire, 1992, Tulving, 1972, 1983, 1986, 1993b).

1.12 *Direct and indirect tests of memory*

The research emphasis on explicit and implicit memory has been driven by findings of dissociations in performance on two types of memory task which differ in the instructions that are given to subjects at the time of testing. On *direct* memory tasks subjects report on some aspect of a prior episode, such as the presence/absence of a word on a previously studied list. By contrast, *indirect* tests make no reference to a prior episode. In this case influences of prior exposure are inferred by analysing the change in some relevant aspect of behaviour, such as an increased probability of correctly identifying a stimulus on a perceptual identification task if that word has been presented in a prior study phase (Jacoby and Dallas, 1981; Neisser, 1954).

Perhaps the strongest evidence for the implicit/explicit distinction stems from studies of the memory deficits of amnesic patients: profoundly amnesic subjects can show little or no evidence of memory for a prior study episode on some direct memory tests, but when tested indirectly they exhibit reliable influences of prior exposure which can be of equivalent magnitude to those observed in normal subjects (Gardner, Boller, Moreines and Butters, 1973; Graf and Schacter, 1985; Graf *et al.*, 1984; Jacoby and Witherspoon, 1982; Shimamura and Squire, 1984; for important precursors see Warrington and Weiskrantz, 1968, 1970, 1974). These findings are consistent with the view that at least some of the processes which support performance on direct tests of memory are not necessary for performance on indirect tests.

Dissociations in performance on direct and indirect memory tasks have also been revealed in studies with normal subjects (Graf, Mandler and Haden, 1982; Jacoby, 1983; Jacoby and Dallas, 1981; Weldon and Roediger, 1987). For example, Jacoby and Dallas (1981) compared performance on tests of perceptual identification and of recognition memory for words which had previously been studied in either a semantic or an orthographic encoding condition. Whilst performance on the perceptual identification task was independent of study task, recognition memory performance was better for items encoded semantically. Findings of this form are again consistent with the view that not entirely the same processes contribute to performance on direct and indirect tests of memory. However, the relationship between the processes which contribute to performance on direct and indirect tasks, and the implications of functional dissociations for the neural substrate(s) of implicit and explicit memory are points of vigorous debate (for reviews and opposing views see Roediger, Weldon and Challis, 1989; Schacter, 1987; Squire, 1992; Tulving, 1983).

1.13 *Memory tasks and memory processes*

One important issue which arises out of the previous discussion of performance on direct and indirect memory tasks is the extent to which these types of task exclusively reflect explicit and implicit memory respectively, since it is widely recognised that the processes engaged on direct and indirect tests may not be wholly distinct (Dunn and

Kirsner, 1989; Jacoby and Kelley, 1992). The most commonly cited example of this form is the fact that processes supporting explicit memory may on occasion contribute to performance on indirect memory tests, in particular if subjects become aware of the relationship between the test task and the relevant prior study phase (Bowers and Schacter, 1990; Schacter, Bowers and Booker, 1989).

Of course, it is equally plausible to suggest that processes supporting implicit memory also contribute to performance on direct tests of memory. An example of this form is the dual-process theory of recognition memory due to Jacoby and colleagues (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992). On standard tests of recognition memory subjects make old/new judgements to a list of words, a proportion of which have been encountered in a prior study phase. Jacoby and colleagues propose that one basis for old judgements is supported in part by the same processes which contribute to performance on some tests of priming (Jacoby and Dallas, 1981). This view is contested by proponents of what can be termed the declarative memory view (Cohen and Squire, 1980; Squire, 1982a). By this view, the processes which support declarative (explicit) memory do not overlap with those which support non-declarative (implicit) memory (Squire, 1992; Tulving, Schacter and Stark, 1982).

The different claims made by these two models are one of the central issues addressed in this thesis, which investigates the processes (both neural and functional) which contribute to explicit memory for an event, and for salient details of the context in which the event occurred. These issues are investigated in a series of studies in which subjects performed modified tests of recognition memory, where in addition to making old/new recognition judgements, subjects also discriminate between items which were presented in one of two contexts in a prior study phase. For example, half of the studied items may have been presented auditorily, and half presented visually.

The dual-process model due to Jacoby and colleagues makes predictions about the relationship between memory for prior occurrence, and memory for details of the context of that occurrence, as does a model which incorporates the declarative memory view, and is principally associated with the work of Squire (Squire, 1994; Squire and Knowlton, 1994), and Moscovitch (1992, 1994). The common ground for these competing models is that both propose qualitatively different processes which contribute to memory for an event which is either accompanied or unaccompanied by retrieval of contextual information about that event. The following sections introduce the details of these alternative models, and review the evidence which informs on the characteristics of the processes that contribute to memory for prior occurrence, and memory for study context.

1.2 *Dual-process theories of recognition memory*

The view that different processes support memory for events and for the context of those events is implicit in dual-process theories of recognition memory (Atkinson and Juola, 1974; Humphreys, Bain and Pike, 1989; Jacoby and Dallas, 1981; Mandler, 1980). Prior to the mid-seventies the predominant view was that performance on tests of recognition memory could be modeled along a single dimension, which corresponded to an assessment of a feeling of familiarity, or of memory strength (Brown, 1975). Whilst the particular mechanisms which were claimed to support old/new discriminations vary considerably between models (for example, see Anderson and Bower, 1972; Kintsch, 1967, 1970), the common ground for these models was that recognition memory performance could be described in terms of a single process.

This unitary framework is well suited to a signal detection analysis of recognition memory performance (Green and Swets, 1966; Norman and Wickelgren, 1969), where old and new items are assumed to form two separate but overlapping distributions of memory strength or familiarity. Under a signal detection model, a criterion level of strength is assumed to be set such that items falling above the criterion are judged old, whereas those falling below the criterion are judged new. Recognition memory performance therefore depends upon the criterion which is set, and the degree of overlap between the strength distributions for old and new items (see also Snodgrass and Corwin, 1988). Necessarily, if recognition memory performance is to be above chance levels then the mean level of familiarity/strength for old items must be higher than that for new items.

In opposition to these unitary models of recognition memory, dual-process models propose that, in addition to an assessment of strength or familiarity, recognition judgements can be made on the basis of retrieval of the prior occurrence of the test item - the learning episode (Atkinson and Juola, 1974; Humphreys *et al.*, 1989; Jacoby and Dallas, 1981; Jacoby and Kelley, 1992; Mandler, 1980). These models distinguish between processes supporting recognition with and without retrieval of context, since both familiarity and retrieval of the learning episode support recognition memory judgements, but only the latter supports context judgements.

The most widely employed and fully articulated dual-process model of recognition memory is due to Jacoby and colleagues (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992; Jacoby, Kelley and Dywan, 1989). They propose that recognition memory performance is supported by the independent processes of *recollection* and *familiarity*. Recollection refers to memory for a prior study episode, whereas familiarity refers to a general feeling that a test item has been previously encountered, in the absence of retrieval of the prior episode.

1.21 *Recollection as a basis for recognition judgements*

The fact that recollection - defined as retrieval of a learning episode - can in principle support recognition memory judgements is uncontroversial. However, the extent to which recollection does in fact contribute to judgements on a standard recognition memory test is not entirely clear, since in the typical recognition memory paradigm subjects simply make old/new discriminations. Thus there is no direct means of assessing the basis for task judgements.

One line of evidence for the influence of recollection on standard recognition memory tasks stems from findings that recognition performance is improved when subjects process the meaning of study items (e.g. Jacoby and Dallas, 1981). Comparable effects of the processing of meaning are also found on tests of free recall (Hyde and Jenkins, 1973). If it is assumed that free recall depends upon retrieval of the learning episode, then these findings are consistent with the view that retrieval of the learning episode also contributes to performance on tests of recognition memory.

A second line of evidence for the influence of recollection on standard tests of recognition memory stems from a reaction time study by Hintzman and Curran (1994), who investigated the time course of the processes contributing to recognition memory judgements. They employed the response-to-signal procedure (Reed, 1973; Wickelgren, 1975), in which presentation of test stimuli is followed at variable lags by a response cue. Subjects are required to respond immediately when the response cue is presented, and the procedure therefore controls the processing time permitted before a judgement is required. Comparison of the memory performance of subjects at different lags can then be employed to assess the time following stimulus presentation at which different types or different amounts of information are available for test judgements.

In two experiments Hintzman and Curran (1994; see also Dosher, 1984) plotted the probability of an incorrect 'old' judgement to a new item (a false alarm) as a function of lag for test items which were highly similar, but not identical to, studied items. The resulting plot was an inverted 'U' function, in which the probability of a false alarm initially increased, and then decreased at longer lags. The authors argued that the initial increase in the probability of a false alarm was due to the fact that at short lags responses were made on an assessment of familiarity, which was not sufficient to discriminate between old items and highly similar new items. The latter half of the inverted U function, where the probability of a false alarm decreased with increasing lag, was interpreted as representing item specific recall, which was sufficient to discriminate between old and highly similar new items. This interpretation is also consistent with findings that accurate old/new recognition judgements can be made at reliably shorter lags than can accurate context judgements (Johnson, Kounios and Reeder, 1994). For broadly similar conclusions regarding the temporal relationship between familiarity and recollection see Atkinson and Juola (1968, 1974), but see also Mulligan and Hirshman (1995).

1.22 Familiarity as a basis for recognition judgements

Jacoby and colleagues propose that the experience of familiarity is related to how fluently test items are processed, *and* to whether that fluent processing is attributed to the fact that the item was encountered at study. Familiarity can in principle be employed as a basis for recognition memory judgements only if presentation of an item at study and at test can result in that item being processed more fluently than an item which is presented at test for the first time.

It is this aspect of the definition of familiarity that links this dual-process theory of recognition memory to processes which contribute to implicit memory. On indirect tests such as lexical decision (Scarborough *et al.*, 1979), or perceptual identification (Jacoby and Dallas, 1981), item repetition results in faster reaction times and a higher probability of correct identification respectively, as has previously been noted. These findings are examples of priming (Cofer, 1967), which is summarily defined as an improved ability to detect or process repeated stimuli (Shimamura, 1986; Squire and Zola-Morgan, 1988).

There are a number of possible relations between the processes which support priming and those that support familiarity-based recognition (Mayes, 1992). However, in general, support for the view that a process akin to priming contributes to performance on tests of recognition memory would stem from demonstrations that manipulating the fluency with which test items are processed results in an increased probability of judging that item old. Note that in order to avoid confusing processes with the phenomenal experiences arising from such processes, the process of familiarity as defined by Jacoby and colleagues will hereafter be referred to as *fluency-based recognition*, or simply *fluency*.

1.221 *Relative fluency*

Implicit in the definition of fluency given above is the notion of a comparative process, since fluency-based recognition judgements are necessarily made relative to some reference or baseline value. One proposal is that this comparison occurs between the extra-experimental (baseline) fluency, and the intra-experimental (local) fluency associated with test items (Mandler, Goodman and Wilkes-Gibbs, 1982; Rugg, 1990). By this view, higher values of local fluency, in comparison to baseline, can result from the presentation of items at study and at test, and therefore provide a basis for recognition judgements. Note that these formulations treat items within a test sequence individually, with no focus on inter-item effects. However, it is equally plausible to suggest that a comparative relation exists between levels of fluency across items within a given test list (Gillund and Shiffrin, 1984).

1.222 *Evidence for fluency-based recognition*

As previously noted, the strongest evidence for the influence of fluent perception on recognition memory judgements would stem from studies which directly manipulated the clarity with which recognition memory test items were presented. Whittlesea, Jacoby and Girard (1990) presented visual recognition memory test stimuli which were occluded by either a heavy or a light visual mask. An equal number of old and new test items were associated with each level of mask, and subjects were not informed of the masking manipulation. For both old and new test items the probability of an old judgement was higher for those items presented with the light mask, suggesting that fluent perception of items does in fact influence recognition memory judgements.

A similar result was obtained by Jacoby and Whitehouse (1989), who briefly presented test items below the threshold of conscious perception immediately prior to the presentation of the same item for a recognition memory judgement. Compared to test items which had not been preceded by a brief stimulus presentation, those that had were associated with an increased probability of a subsequent old judgement, regardless of whether test items were in fact old or new.

These findings support the view that fluent perception can serve as a basis for recognition memory judgements if it is assumed that the subliminal presentation of test items increased the fluency with which the test items were perceived on the subsequent presentation. The findings in these two studies complement a series of studies which have reported correlations between measures of fluent perception and recognition memory performance (Johnston, Dark and Jacoby, 1985; Johnston, Hawley and Elliot, 1991; Kelley, Jacoby and Hollingshead, 1989). These manipulations have focused wholly on the perceptual features of test stimuli. However, Whittlesea (1993) has demonstrated that task manipulations which influence the processing of the meaning of test items may also influence recognition performance. For example, in a modified recognition test subjects made recognition judgements to the final words in a sentence, where words were presented one at a time to the subject (Whittlesea, 1993, experiment 2). The final word was either predictable on the basis of the preceding words ('The man withdrew his money from the *bank*'), or was relatively ambiguous ('The man withdrew his money from the *shoebox*'). The probability of an 'old' judgement to final words which were presented in the predictable context was greater than that in the ambiguous context, irrespective of the actual old/new status of the test item. If the predictable context increased the fluency with which subjects processed the meaning of the words, then these findings are consistent with the view that the basis for fluency judgements is not restricted to the domain of perceptual processing.

1.223 *Attributed fluency*

As previously noted, fluency-based recognition is based upon fluent processing *and* a subsequent attribution, where the increased fluency is attributed to the fact that the item has been encountered previously. Evidence for this attribution process comes from the same experiments which demonstrated the influence of fluent perception on recognition memory judgements. For instance, in the study of Whittlesea, Jacoby and Girard (1990), the effect of masking on recognition memory performance was eliminated when subjects were informed that the clarity with which test items were presented would be varied. Similarly, when Jacoby and Whitehouse (1989) increased the duration of the brief stimulus presentation so that subjects were aware of this presentation, there was no subsequent increase in the probability of 'old' recognition memory judgements.

Considered jointly, these results support the claim that item memory judgements are not based on fluent perception *per se*. Rather they are made on the basis of an attribution, where under certain circumstances the fluency with which a test item is processed will be attributed to the fact that the item has been encountered before. Jacoby (1992) proposes that the fluency with which items are processed will be attributed to whichever source is the most likely given the options available in a particular task (see also Jacoby, Allan, Collins and Larwill, 1988; Jacoby, Woloshyn and Kelley, 1989).

1.3 *The process-dissociation procedure*

On a standard old/new test of recognition memory the processes of fluency and recollection both support correct old judgements. However, it is not possible to estimate the contributions that each process makes to task performance, because the contribution of each results in the same behavioural outcome - an old response. Jacoby and colleagues (Jacoby, 1991; Jacoby, Toth and Yonelinas, 1993) have introduced the process-dissociation procedure (hereafter PDP) as a means of estimating the respective contributions of recollection and fluency to task performance.

In order to obtain these estimates the PDP compares task performance under so-called inclusion and exclusion instructions. The estimates are obtained by applying mathematical formulae to describe performance on the two tasks, yielding an inclusion score and an exclusion score. In the study phase for both tasks subjects encounter items presented in one of two contexts. So, for example, half of the items may be presented auditorily, and half visually. Under inclusion test instructions subjects simply make old/new judgements to test items, regardless of the context in which the items were presented at study. Under exclusion instructions subjects make an old response only to items presented in one of the two contexts (targets), and respond new to items presented in the alternate context (non-targets), as well as to genuinely new items.

Under inclusion instructions both recollection and fluency support correct old judgements. The inclusion score is therefore the probability that an item is recollected $p(R)$, plus the probability that an item is judged old on the basis of relative fluency $p(F)$, minus the intersect of these two $p(F \cap R)$

$$\text{inclusion score} = p(R) + p(F) - p(F \cap R) \qquad \textbf{(1)}$$

The exclusion score is defined as the probability of an incorrect old judgement to a non-target. It is assumed that an incorrect old response to a non-target must have been made on the basis of fluency, since recollection would have permitted a correct new judgement. The exclusion score is therefore defined as the probability that an item is associated with fluency-based recognition p(F), less the probability that the item is associated with fluency and recollection $p(F \cap R)$.

$$\text{exclusion score} = p(F) - p(F \cap R) \qquad \textbf{(2)}$$

Note that this description of performance under exclusion instructions entails the strong assumption that when a subject is aware that an item is old, but is not in possession of sufficient information to make a context discrimination, then the subject will invariably respond old. Task instructions have, with one recent exception (Yonelinas and Jacoby, 1995), not informed subjects how to respond when an item is familiar but the context is not known. This issue was addressed recently by Rugg, Allan and Wilding (1995), who compared performance on two exclusion tasks in which the instructions given to subjects at test differed. On the first, subjects were instructed always to respond new when an item was familiar but not recollected. On the second, subjects were instructed always to respond old to familiar but unrecollected items. The exclusion score in the latter condition was higher, suggesting that unless subjects are explicitly instructed how to respond in this circumstance it is uncertain what strategy they will adopt. If this is true, then the accuracy of the exclusion score obtained is questionable. Other considerations related to this point will be discussed in chapter 9, on the basis of electrophysiological data obtained while subjects performed a recognition memory exclusion task.

1.31 *Estimating recollection*

Having obtained inclusion and exclusion scores from equations (1) and (2) above, it is then possible to obtain an estimate of p(R) by subtracting equation (2) from equation (1):

$$\text{inclusion} - \text{exclusion} = p(R) \qquad \textbf{(3)}$$

For this procedure to be valid it is necessary to assume that the estimates of R and F obtained under inclusion and exclusion instructions are equivalent. The comments made above regarding task instructions and the exclusion score suggest that this may not hold for fluency-based recognition. The assumption of invariance of recollection has also been questioned, principally on the grounds that the assumption requires that recollection is equally probable (or equivalently engaged) when it is necessary for the task discrimination (exclusion task) and when it is incidental to that discrimination (inclusion task) (Graf and Komatsu, 1994; Jacoby *et al.*, 1993; Yonelinas and Jacoby, 1994).

However, as Toth and colleagues observe (Toth, Reingold and Jacoby, 1995), the proposal that recollection is not invariant across inclusion and exclusion tasks on tests of recognition memory is based on theoretical and not empirical grounds. In addition, recent work has proposed a means of obtaining inclusion and exclusion scores from a single exclusion task (Yonelinas, 1994; Yonelinas and Jacoby, 1994). Under this model the exclusion score is obtained as previously described, and the inclusion score is defined as the probability of a correct old judgement to a target. In principle this formulation is not susceptible to the criticisms outlined above, since the estimates of R and F are obtained from the same task. If this modified version of the PDP is to yield reliable estimates, then R and F must either be the same for targets and for non-targets, or subjects must perform two exclusion tasks in which the target/non-target distinction is reversed for the particular context discrimination in which subjects engage.

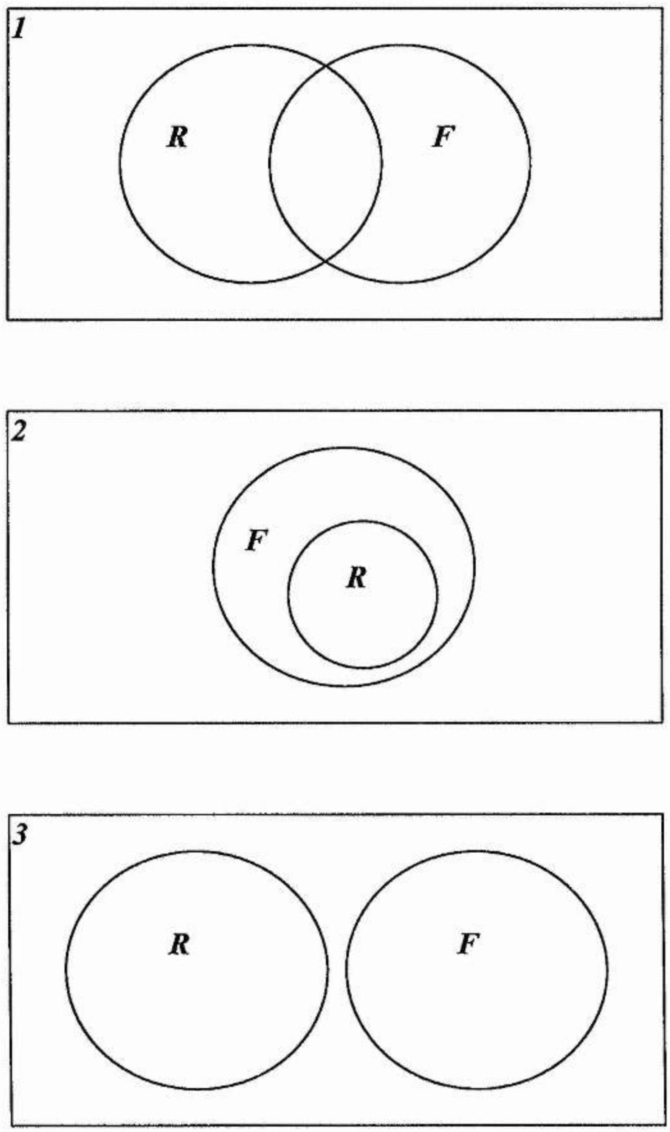
1.32 *Estimating fluency*

Having gained an estimate of R from equation (3) it is necessary to make a further assumption regarding the relationship between R and F in order to obtain an estimate for fluency, since the term $p(F \cap R)$ can only be computed if the nature of the intersection between R and F is explicitly specified. Three possible relations will be described here, and they are all displayed in diagram form in figure 1.1 overleaf (see also Jones, 1987).

The upper panel of figure 1.1 displays a relationship of independence between R and F. This is the relationship proposed by Jacoby and colleagues (Jacoby, 1991; Jacoby *et al.*, 1993; Jacoby, Toth, Yonelinas and Debner, 1994; Jacoby, Yonelinas and Jennings, in press). Under this model a test item may be judged old on the basis of recollection, on the basis of fluency, or be associated with both recollection and fluency. Under an independence assumption the intersection of R and F is simply the product of these two factors. Therefore equation (2) becomes:

$$\text{exclusion} = p(F) - p(F)p(R) = p(F)(1 - R) \qquad \textbf{(4)}$$

Figure 1.1. Alternative models for the relationship between the processes of Recollection (R), and Fluency (F). Panels 1, 2, and 3 describe relationships of Independence, Redundancy, and Exclusivity respectively.



Given a value for R equation (4) can be solved for F.

Two alternative relationships between R and F are also depicted in figure 1.1. The middle panel denotes a relationship of redundancy, whereby the set of recollected items is a subset of the set of items that would be judged old on the basis of fluency (Joordens and Merikle, 1993). By this model, $p(F \cap R)$ is simply equal to R. Therefore equation (2) becomes:

$$\text{exclusion} = p(F) - p(R) \qquad \textbf{(5)}$$

Finally, the third panel of figure 1.1 depicts a relationship of exclusivity between R and F. By this view an item is either recollected, or judged old on the basis of fluency. Under this model $p(F \cap R) = 0$, thus F is simply equal to the exclusion score.

Support for the view that independence is the best characterisation of the relationship between recollection and fluency has been claimed on the basis of findings that estimates gained from the PDP when this relationship is assumed are consistent with experimental findings on indirect and direct memory tests. For example, dividing attention at study had no effect on the estimate of fluency, but resulted in an attenuation of the estimated probability of recollection (Jacoby and Kelley, 1992). However, the logic of this argument is circular, since a model assuming independence has been employed to support the independence assumption.

Stronger support for the independence assumption stems from the lack of fit between experimental data and the assumptions underlying the redundancy and exclusivity models (although see Curran and Hintzman, 1995; Russo and Andrade, 1995). The divided attention manipulation cited above is a case in point. In the study reported by Jacoby and Kelley (1992), divided attention resulted in poorer overall recognition performance. However, there is evidence that priming is unaffected by the divided attention manipulation. For example, Parkin and colleagues (Parkin, Reid and Russo, 1990) employed a divided attention manipulation at study and compared subsequent performance on tests of recognition memory and word fragment completion - an indirect memory test where subjects are instructed to complete word fragments (e.g. _ss_ss_n) with a suitable word (assassin). Recognition performance was lower in the divided attention condition, but word fragment completion performance was independent of the study manipulation.

To the extent that equivalent performance under conditions of full and divided attention reflects intact priming, these findings are problematic for a redundancy model, which predicts that all recollected items would also have been judged old on the basis of fluency. Consequently, under a model of redundancy it should not be possible to find experimental manipulations which have no influence on fluency and result in poorer overall memory performance. Evidence against the exclusivity model of recollection and fluency will be returned to after the introduction of an alternative means of assessing the processes which contribute to performance on tests of recognition memory.

1.4 *The remember/know procedure*

The Remember/Know procedure, initially introduced by Tulving (1985), directly addresses the issue of the states of awareness of subjects whilst performing recognition memory tasks. In this experiential approach subjects report at test whether their recognition memory judgements are based upon a 'recollective experience' (a 'Remember' or 'R' response), or on a general feeling that a test item was encountered at study, in the absence of a recollective experience (a 'know' or 'K' response). The procedure is therefore an alternative means of investigating the processes that contribute to performance on recognition memory tasks.

The R/K paradigm has been most extensively employed by Gardiner and colleagues (for a review see Gardiner and Java, 1993). Their findings in a number of studies are similar to those obtained using the PDP to estimate the contributions of R and F to recognition memory performance. For example, study manipulations such as depth of encoding or dividing attention have been shown to selectively influence the probability of an 'R' response, and the probability that a word is recollected as estimated by the PDP (Gardiner, 1988; Gardiner and Parkin, 1990; Jacoby and Kelley, 1992). There is also evidence that manipulations of the fluency with which test items are processed selectively influence the probability of a 'K' response (Rajaram, 1993).

However, a critical disparity between the views of Gardiner and colleagues and those of Jacoby concerns the relationship between the processes of recollection and fluency. Gardiner and Java (1993) propose that at the level of conscious experience the relationship between 'R' and 'K' responses is one of exclusivity. However, if the R/K procedure is to be employed to make inferences about the processes that support recognition memory judgements then it is necessary to assume a relationship at the level of the processes themselves. Consequently, Gardiner and colleagues implicitly assume a relationship of exclusivity between the processes of recollection and fluency, where the probability of a K response equals the probability of a response based on fluency.

Yonelinas and Jacoby (1995) directly compared the estimates of R and F gained from the PDP under an independence assumption with values of R and F obtained from the R/K procedure. In the recognition memory test that they employed the test items were either size congruent or size incongruent (smaller or larger) relative to the study items. Using the PDP, p(F) was lower for size incongruent objects, whilst p(F) obtained using the R/K procedure was higher for incongruent objects. Yonelinas *et al.*. (1995) argued that the value of F obtained by the R/K procedure was inconsistent with the view that perceptual similarity should increase the probability of a response made on the basis of familiarity. They suggested that the disparity between the findings using the PDP and the R/K paradigm arose because the use of uncorrected R/K data assumes a relationship of exclusivity. They demonstrated that if a relationship of independence is assumed to hold between remember and know responses then the two procedures provide equivalent estimates of R and F. This demonstration was achieved by dividing the proportion of K responses by the proportion of responses to which an R response was not made. With reference to the equations outlined above, this is the equivalent of setting the probability of a K response equal to the exclusion score in equation (4), and solving the equation for F.

Using the R/K paradigm, Knowlton and Squire (1995, experiment 3) recently reported further findings which are difficult to reconcile with the view that exclusivity describes the relationship between R and K responses. The authors compared the probability of R and K responses at lags of 10 minutes and 1 week after an initial study phase. A significant proportion of items which had attracted an R response at the short delay attracted a K response at 1 week. This finding is difficult to reconcile with an exclusivity relationship, since the two bases for judgements are assumed to be entirely distinct, and consequently there should be a negligible rate of conversion from R to K responses. The pattern of results reported by Knowlton and Squire (1995) is consistent with either a redundancy or an independence model of the relationship between R and K responses, since both relationships are consistent with the view that some initially recollected items will be assigned a K response at the longer delay (Knowlton and Squire, 1995). When considered along with the evidence discussed previously that divided attention at study attenuates overall recognition memory performance, the available evidence therefore supports the view that the independence model best characterises the relationship between recollection and fluency.

However, the results of Knowlton and Squire (1995) are also consistent with the view that the R/K distinction simply reflects some form of graded recollection. At the longer delay overall memory was poorer, and the probability of an R response decreased while the probability of a K response remained the same at short and long delays. These findings are wholly consistent with the view that subjects maintained the same criteria for distinguishing R and K responses at the two lags, and that memory for events was weaker at the longer delay, thereby resulting in a diminution in the proportion of R responses, and an increase in the number of forgotten items. There have been no PDP studies to date which have investigated rate of forgetting on the estimates of recollection and fluency.

The concept of graded recollection, introduced above, has not received a great deal of attention, and, to anticipate, it is one that will be returned to when considering the electrophysiological data reported in the empirical sections of thesis. At this juncture it is worth noting that Jacoby and colleagues (Jacoby *et al.*, 1993; Yonelinas, 1994) argue that recollection is an all-or-none process, and invariably results in a highly confident old judgement on recognition memory tests. This concept was employed by Yonelinas (1994), who investigated the distribution of confidence judgements elicited while subjects performed a recognition memory exclusion task. In three experiments subjects were required to make memory judgements to old and new test items, and in addition to rate their confidence in each judgement on a six-point scale, which ranged from '*sure the item is old*' to '*sure the item is new*'. The context discrimination that subjects were required to make on the exclusion tasks was between items that had been presented in one of two lists in a prior study phase. The analyses performed on these data revealed a skewed distribution of responses, wherein a disproportionately high number of highly confident and correct judgements were made to old items.

Yonelinas (1994) argued that these findings were inconsistent with a unitary process model of recognition memory, since if old and new items were distributed along a linear continuum of familiarity or memory strength, and if confidence judgements were directly related to the strength of familiarity, then the distribution of confidence judgements should be equivalent for old and new items. Yonelinas (1994) interpreted these contradictory findings within the dual-process framework of Jacoby and colleagues (Jacoby and Dallas, 1981), suggesting that the skewed distribution for old items reflected the influence of a second process on recognition decisions. He proposed that this second process was recollection, which, when engaged, resulted in a correct and highly confident old judgement (see also Jacoby *et al.*, 1993).

The applicability of these findings to tests of recognition memory in which subjects do not make an explicit context judgement can be questioned, since it is not clear that the distribution of confidence judgements would be comparable if subjects were required simply to distinguish between old and new test items. Further, in one sense the definition of recollection as an all-or-none process which results in a highly confident response is necessary if the PDP is to be employed to accurately estimate recollection and fluency. If graded recollection was sufficient in some circumstances to support old judgements but not context judgements, then there would be no means of separating responses of this type from those made on the basis of fluency. This is an important point, and it emphasises that results from the PDP approach do not constitute evidence that recollection and fluency contribute to recognition memory performance: it is necessary to assume that they do so before employing the relevant equations.

In one study in which this error is made (Verfaellie and Treadwell, 1993), estimates of R and F were obtained for an amnesic subject group and a group of matched controls. For words that had been heard in a prior study phase the PDP equations estimated probabilities of R and F of 0.00 and 0.46 for the amnesic patients, and values of 0.33 (R) and 0.50 (F) for the controls. These findings were interpreted as support for the view that the process of fluency contributes to recognition judgements and is unimpaired in amnesia. However, the findings only permit the conclusion that amnesics and controls were, on a proportion of test trials, in possession of sufficient information to make an old judgement, but insufficient information to make a correct context judgement.

In addition, two further aspects of the findings in the study of Verfaellie and Treadwell (1993) make interpretation of the data problematic (Roediger, 1994). First, for items that had been read at study the PDP estimates of R and F were equivalent for amnesic and control subjects. It is unclear why performance for amnesic and control subjects should differ so markedly for test items which were heard, whilst no such disparity is evident for items that were seen. Second, and more fundamentally, the probability of a false alarm - an incorrect old judgement to a new test item - was markedly higher for the amnesic subject group. In the absence of an appropriate correction for this disparity, an interpretation of the findings of Verfaellie and Treadwell (1993) is not straightforward (but see Verfaellie, 1994).

A correction for false alarms in the PDP framework has recently been proposed (Jacoby *et al.*, in press), which assumes that false alarms are typically made on the basis of fluency. Correspondingly, the correction is applied only to the estimate of fluency gained from PDP equations. An alternative view is that when false alarm rates differ across subject groups, a comparison of any differences between the estimates gained from the PDP equations will not yield results which are interpretable (Roediger, 1994; see also Buchner, Erdfelder and Vaterrodt-Plünnecke, 1995).

1.5 *Declarative and non-declarative memory*

An alternative to the dual-process theory view of the relationship between recognition without and without retrieval of context is that both these forms of memory depend exclusively upon what has been termed the declarative memory system (Cohen and Squire, 1980; Squire and Knowlton, 1994). This system is proposed to underlie all forms of memory retrieval which are accompanied by the experience of remembering. Declarative memory is contrasted with non-declarative memory, which refers to a variety of forms of memory which manifest themselves in behaviour in the absence of an accompanying memory experience. The distinction between declarative and non-declarative memory is therefore very similar to that between explicit and implicit memory (Graf and Schacter, 1985).

By the declarative memory view, recognition memory cannot be influenced by relative fluency, since priming is a form of non-declarative memory (Squire, Shimamura and Graf, 1985; Squire and Zola-Morgan, 1988; Tulving, 1983; Tulving *et al.*, 1982). One line of experimental evidence relevant to this proposal comes from studies which have compared the recall and recognition memory performance of normal and amnesic patients. The rationale for these studies was that patients with amnesia are severely impaired on direct tests of memory, including recognition and recall. However, on indirect tests of memory, including studies measuring priming, their performance is often equivalent to that of unimpaired subjects, as has previously been noted (Richardson-Klavehn and Bjork, 1988; Shimamura, 1986). If recognition memory, but not recall, in part depends upon processes related to priming, then relative to the recall and recognition performance of normal subjects, amnesics should show some savings on tests of recognition compared to their performance on tests of recall.

In two studies Hirst and colleagues (Hirst, Johnson, Phelps, Risse and Volpe, 1986; Hirst, Johnson, Phelps and Volpe, 1988) found precisely this pattern of results. These findings are consistent with the view that a second process, intact in amnesic patients, contributes to recognition memory judgements. However, in contrast to these findings, Haist, Shimamura, and Squire (1992) failed to replicate the findings of the second experiment by Hirst and colleagues (1988). Haist *et al.*. (1992) also compared recognition and recall performance of amnesic patients over different time periods, and elicited confidence judgements in the recognition decisions. They found no evidence for a relative sparing of recognition memory performance in the amnesic patients. The confidence judgements were employed to assess the strength of feelings of familiarity which accompanied recognition judgements. These confidence judgements were found to be predictive of performance on the recall task, suggesting that the same processes contributed to both types of judgement.

One suggestion for the disparate findings across these studies is that in the second study of Hirst and colleagues (Hirst *et al.*, 1988) some of the amnesic patients also had additional damage to frontal brain systems, which may have resulted in a decrement in their performance on the recall task which was proportionately greater than the associated decrement on the recognition task (Squire *et al.*, 1993). This view is supported by findings that damage to the frontal lobes can affect recall performance more than recognition (Jetter, Poser, Freeman and Markowitsch, 1986).

A recent study of the memory deficits following ECT induced amnesia has also investigated the relationship between priming and recognition memory (Dorfman, Kihlstrom, Cork and Misiaszek, 1995). ECT (electro-convulsive therapy) has been shown to impair performance on direct tests of memory, while sparing performance on certain indirect tests (Graf *et al.*, 1984; Squire *et al.*, 1985). In the study of Dorfman *et al.*. (1995) recognition performance was assessed on two recognition memory tasks, and priming was assessed on a test of stem completion, where subjects were presented with a three letter stem and instructed to complete the stem with the first word that came to mind. Priming is measured on stem completion by the above baseline probability of completing a stem with a previously studied word. On this completion task both amnesics and controls showed reliable evidence of priming.

In the first recognition memory task subjects were told to respond old only if highly confident in the test judgement, while in the second task subjects were told to respond old even if uncertain of the status of a test word. Recognition memory performance was only reliably greater than chance in the second task. The authors assumed that the improved memory performance in the second task resulted from an increased probability of an old response made on the basis of a process akin to fluency. This interpretation rests on the assumption that responses based on fluency are typically relatively low confidence recognition judgements. In addition, it is not clear why the findings cannot be explained by assuming that subjects were in possession of weak or residual recollection, which was revealed only under the less stringent criteria for old judgements set in the second recognition task.

Knowlton and Squire (1995, experiment 1) investigated the bases for recognition memory judgements in amnesic and control subjects using the R/K procedure. Compared to controls, the amnesic subjects showed a diminution in the probabilities of R and of K responses. These findings are inconsistent with the view that a process akin to priming contributes to recognition judgements, since if this were the case then the probability of a K judgement should have been equivalent for amnesics and controls. Rather, the results are consistent with the view that R and K judgements are both related to the processes which support declarative memory.

Another approach to investigating the relationship between the processes supporting recognition memory judgements is to assess the degree of dependence or independence between performance on direct and indirect memory tests (Tulving *et al.*, 1982). According to the declarative memory hypothesis, performance on direct and indirect tasks should be independent. However, if the processes which contribute to priming also contribute to recognition judgements, then under some circumstances a relationship of dependence should hold between performance on direct and indirect tasks (Kelley *et al.*, 1989). The available evidence is mixed: on some tasks a relationship of independence has been reported (Jacoby and Witherspoon, 1982; Tulving *et al.*, 1982), and on others a relationship of dependence has been revealed (Jacoby and Witherspoon, 1982; Shimamura and Squire, 1984).

The validity of this approach has also been questioned, the most telling criticisms centering on the fact that a demonstration of stochastic independence requires that the same test items are presented to the same subjects on both direct and indirect tasks (Richardson-Klavehn and Bjork, 1988; Shimamura, 1985). The principal problem with this approach has been well articulated by Richardson-Klavehn and Bjork (1988, page 497): "It is simply an implausible assumption that the item exposures and subject reactions that comprise a first test leave the memory system unaltered and ready to give an uncontaminated picture of the influence of the study episode on a second test of some type." A second criticism of the procedure employed to measure the relationship between performance on direct and indirect tasks is that the relationship revealed by an analysis of the complete data set may not be reflective of the relationship which holds for significant subsets of the data (Hintzman, 1980). Whilst one method for circumventing these putative problems has been proposed (Hayman and Tulving, 1989), the focus in the contemporary literature has moved away from this potential source of evidence for the relationship between the processes which contribute to performance on direct and indirect tests of memory.

In conclusion, the extent to which priming contributes to recognition memory in amnesic patients is unclear, contrasting with the relatively strong evidence for fluency-based recognition in normal subjects (Jacoby and Kelley, 1992; Whittlesea, 1993). Squire has suggested that fluent processing may influence recognition memory judgements only when overall memory for test items is poor (Squire *et al.*, 1993; Squire and Knowlton, 1994), and a similar proposal has been advanced by Johnston (1991). A second explanation for the inability of amnesic patients to use attributed fluency is that these patients have no explicit memory for the study episode, and therefore cannot make an attribution to that episode (Jacoby and Kelley, 1992). The precise relationship between fluency-based recognition and the recognition memory performance of amnesic patients is unresolved.

1.51 *Memory for context and the declarative memory system*

A consequence of the view that recognition with and without retrieval of context both rely on declarative memory is that they share a common neural substrate. This is generally agreed to comprise the medial temporal lobes and diencephalic structures, damage to which can lead to selective and severe deficits of all forms of declarative (explicit) memory (Milner, 1966; Scoville and Milner, 1957; for early animal models see Gaffan, 1974; Hirsh, 1974; for reviews see Mayes, 1988; Squire and Zola-Morgan, 1991; Zola-Morgan and Squire, 1993). By contrast, selective damage to the frontal lobes reveals more specific memory deficits (Shimamura, 1994; Stuss, Eskes and Foster, 1994). In particular, frontal lobe pathology has been associated with poor memory for a variety of contextual details of prior learning episodes, such as where or when specific information was acquired, or who provided a particular item of information (Janowsky, Shimamura and Squire, 1989; Milner, 1971; Schacter, Harbluk and McLachlan, 1984; Shimamura and Squire, 1987; Squire, 1982b).

These findings have led to proposals that the two forms of memory - recognition with or without retrieval of contextual or source information - are both functionally and neurologically dissociable (Moscovitch, 1992; Moscovitch, 1994; Squire and Zola-Morgan, 1988). Direct support for this view in a non-amnesic subject population stems from a recent study investigating memory for sentences, and memory for the voice in which the sentences were spoken (Glisky, Polster and Routhieaux, 1995). In this study, elderly subjects were assessed on tests of frontal lobe and medial temporal lobe function. Performance on the tests of frontal lobe function was independent of memory for sentences. However, those subjects who scored highly on the frontal tests exhibited better memory for the voice in which the sentences had been spoken than those subjects who scored low on the tests of frontal function. By contrast, subjects who scored highly on the tests of medial temporal lobe function exhibited better memory for sentences than those who scored poorly. However, memory for voices was independent of subjects scores on the medial temporal lobe test. This double dissociation between memory for sentences and the context in which they were presented is consistent with the view that recognition with and without retrieval of context depends upon distinct neural systems.

Moscovitch (Moscovitch, 1994) proposes that the output of medial temporal lobe structures is accompanied by a sense of familiarity, which denotes whether an item or event has been previously encountered (see also Milner, 1989). He further suggests that the output from medial temporal lobe structures is subject to further processing by the frontal lobes, the function of which is to situate memories in their appropriate spatio-temporal context. Moscovitch views the frontal-lobes as 'working-with-memory' structures, which operate upon the products of retrieval (Moscovitch, 1994).

A similar view is espoused by Squire and colleagues (Knowlton and Squire, 1995; Squire and Knowlton, 1994; Squire and Zola-Morgan, 1988), who suggest that the medial temporal lobes are sufficient to support the feeling of knowing associated with K judgements in the R/K paradigm, whereas the integrity of the frontal lobes is necessary for recollection of the study episode, as revealed by R judgements (see especially Knowlton and Squire, 1995).

In functional terms these two views propose that recollection consists of a retrieval function and an integrative function. These proposals therefore suggest a relationship of redundancy between recognition with and without retrieval of context, where the probability of recognition will always be equal to or greater than the probability of retrieval of study context.

1.6 *Summary*

Two competing models have been introduced which describe the relationship between recollection with and without retrieval of context. Whilst both models propose that different processes contribute to these two forms of explicit memory, they make very different predictions concerning the nature of, and the relationship between, these processes. The behavioural evidence from normal and amnesic patients does not overwhelmingly favour either model, and in particular the role of fluency on tests of recognition memory is unresolved.

An additional means of deciding between these opposing views is to determine whether neural activity associated with recognition with and without retrieval of context dissociates in a manner more compatible with one or other account. This is one of the principal aims of this thesis, where event-related potentials (ERPs) are employed as the measure of neural activity.

Since ERPs are an on-line measure of neural activity, they can in principle inform on both the nature and the time course of the processes which contribute to memory for a prior episode, and it is well established that ERPs differentiate correctly recognised old and new items on tests of recognition memory (for reviews see Rugg, 1994; Johnson, 1995). However, the functional significance of the memory-related ERP effects which have been reported is not fully resolved. In addition, scant attention has been paid to the sensitivity of ERPs to retrieval of contextual information.

The following chapter introduces the use of ERPs as a research tool, and discusses methodological and conceptual issues regarding the use of ERPs in psychological research. Following this, chapter 3 reviews the relevant studies of memory in which ERPs have been recorded, introduces the general paradigm employed in the empirical chapters comprising this thesis, and outlines the patterns of ERP data which would constitute evidence for either of the competing models described above.

Chapter 2

2 *Event-related potentials: Principles and recording techniques*

As is the case with any measure employed to inform on aspects of behaviour, the ERP technique has a number of limiting factors that constrain both the range of tasks within which it may be usefully employed, and the inferences that may be made on the basis of any results obtained (Hillyard and Kutas, 1983; Hillyard and Picton, 1987; Rugg and Coles, 1994). The following sections discuss the electro-genesis of ERPs, and the ways in which ERP waveforms are described. Additional issues discussed are the constraints that apply and the assumptions that are required when making functional inferences on the basis of this form of electrophysiological data.

2.1 *The electrogenesis of ERPs*

Scalp recorded ERPs are time slices of the ongoing EEG. They represent changes in the electrical activity of the brain which are time-locked to a particular event, such as the presentation of a word on a TV monitor. ERPs recorded at the scalp reflect the summated electrical activity of large populations of cells (Allison, Wood and McCarthy, 1986; Nunez, 1981; Nunez, 1990). At the level of individual cells this electrical activity is due to the bi-directional flow of positive and negative ions resulting from changes in cellular membrane permeability. This activity may in principle be recorded at a distance due to the fact that brain tissue, skull, and scalp are conductive media.

The electrical activity that is in fact recorded at the scalp depends upon the location and structure of cellular configurations, and the temporal characteristics of the electrical activity of cells within and between configurations. Particular configurations of cells generate fields that extend beyond the bounds of the configuration itself. These are termed 'open' fields. An open field configuration is a necessary but not a sufficient condition for distant field recording. A further requirement is that the cells in a given configuration are synchronously active. Hence, a layer of cells oriented in the same direction and in which synchronous activity occurs will generate a summated field that can be recorded at the scalp. Other configurations of cells produce field potentials that may not be recorded outwith the bounds of the configuration. These are termed 'closed' fields, wherein the orientation and temporal characteristics of the cellular elements summate such that the potential field may only be recorded locally (Wood, 1987).

2.2 *ERP recording*

The basic unit of data elicited in ERP recording is a measure of the potential difference between two scalp locations. An ERP waveform is a sequence of such data points, sampled at discrete intervals. The sampling rate of this analogue-digital conversion determines the temporal resolution of the resulting waveform. The rate must be such that it captures all frequencies of interest within it (Cooper, Osselton and Shaw, 1980; Picton, Lins and Scherg, 1995).

Typically, ERPs are concurrently recorded from midline and lateral scalp sites, with locations defined by an accepted system (Jasper, 1958). Recordings at each site are made with respect to a common reference point. Consequently, while the absolute value of the potential difference at any point in the electrical field depends upon the choice of reference, the profile of the field is reference independent (for further relevant comments see Binnie, 1987).

2.3 *Signal extraction*

The waveform resulting from one sample of EEG can be assumed to be composed of two parts (John, Ruchkin and Vidal, 1978). The critical part is the neural activity (the signal) evoked by the particular stimulus in a given task. The second part of the EEG sample is noise. This latter component consists of neural contributions to the ERP waveform that are unrelated to the presented stimulus, as well as non-neural contributions such as muscle activity and eye movements. It is critical that suitable signal extraction procedures are employed, in order to separate the task related and the task unrelated aspects of the EEG.

2.4 *Signal averaging*

The most widely employed signal extraction procedure applied to ERPs is signal averaging, which is performed over the point by point digital values. This is the technique employed in the studies reported in this thesis. The assumption entailed when applying this procedure to ERPs is that the 'noise' in a given sample of EEG is random. Therefore, averaging across trials will reduce the impact of the noise in the averaged ERP, whilst leaving electrophysiological activity which is constant across trials unaffected. The greater the number of trials contributing to the average, the higher the signal/noise ratio.

Signal averaging therefore requires experiments to be designed such that a class of ERPs elicited under precisely the same conditions are produced. Variations in the amplitude, or more critically the latency, of an ERP signature ('latency jitter') across individual trials will result in an unrepresentative average. Similarly, an averaged ERP may in principle reflect an ERP signature which is present on only a proportion of the trials comprising an average. One means of assessing whether averaged ERPs accurately reflect the activity in single trials is to inspect individual samples of the EEG and to measure the amplitude and/or latency of particular peaks or troughs. Whilst this is possible in some paradigms, the low signal/noise ratio for individual trials often precludes this form of analysis.

Another means of increasing the signal/noise ratio is to reject certain classes of trials prior to averaging. A common source of EEG contamination is due to eye blinks and eye movements, both of which cause changes in potential over anterior scalp locations (Lins, Picton, Berg and Scherg, 1993). Concurrent electro-oculargram (EOG) recording permits monitoring of eye blink artifacts. One approach to reducing any artifactual contribution due to blinks is to establish a criterion within which activity on the EOG channel must fall in order for a given trial to be accepted. This method is employed in the experiments reported in this thesis, where subjects are instructed to minimise blinks during the critical recording phase.

The influence on the averaged ERPs of other more general non-neural artifactual contributions such as muscle movement can also be attenuated by setting a criterion within which activity at all scalp-sites must fall, and rejecting trials on which the electrical activity on any recording channel falls outside this criterion. This procedure is valid to the extent that all task-related neural activity falls within the range of the criteria that are set.

2.5 *Descriptions of ERPs*

The changes in the electrical activity of the brain denoted by ERPs can be characterised as a series of peaks and troughs. These changes in potential are classified as 'exogenous' or 'endogenous' depending upon whether they are determined mainly by the form of a stimulus, or the functional effects of the stimulus (Sutton, Braren, Zubin and John, 1965). While it is a general rule of thumb that exogenous components precede endogenous components, the exogenous/endogenous distinction is not dichotomous. Instead, it denotes a continuum, at one end of which are changes in potential which are entirely stimulus bound, and at the other are those which are particularly sensitive to cognitive variables and to task demands (Donchin, Ritter and McCallum, 1978).

Both the exogenous and the endogenous components of an ERP waveform can be described in terms of their latency and polarity relative to a specified reference point. Concurrent recording at a number of scalp sites permits the description of features of the ERP waveform to include variations in latency and polarity as a function of location. These descriptive dimensions are neutral with respect to the causes of the recorded ERP data, and there are a number of ways in principle in which ERPs can be described in terms of their causes. The issue of the optimum descriptive framework is contentious (Picton and Stuss, 1980), but in general two main approaches to the description and classification of ERPs can be identified. The common ground for these approaches is that an ERP waveform can be viewed as a number of components which overlap both spatially and temporally.

2.51 *Physiological component definitions*

Physiological definitions of the components comprising an ERP waveform fall relatively straightforwardly from the description of brain electrophysiology given above. One option is to define a component as the contribution to an ERP field of a single generator process (Näätänen, 1982; Näätänen and Picton, 1987). There is of course no necessary reason why a component should be defined in terms of a single neural source, and correspondingly other solely physiological definitions could in principle be formulated wherein a component is defined in terms of a number of generators or a particular neural circuit.

2.52 *Functional component definitions*

The physiological level of description of ERPs does not require recourse to the conditions under which ERPs are evoked. By contrast, functional approaches to component definition employ the conditions under which different patterns of ERP waveforms are elicited to define components in terms of cognitive processing operations

(Donchin *et al.*, 1978). This approach necessarily involves a comparison of ERPs evoked under different experimental conditions. In this framework, component definitions therefore rely upon inferences concerning the nature of the processes which different experimental manipulations engage.

Given that ERPs are direct reflections of neural activity, functional approaches to component definition are obliged to make some assumptions regarding the relationship between the level of functional description and the level of neural activity. The widely employed framework of Donchin and colleagues (Donchin *et al.*, 1978) assumes that the brain structures underlying an ERP component constitute a functionally distinct unit. There is no requirement for a specification of involved structures, and no requirement that a particular structure contributes solely to one distinct functional unit, meaning that any number of physiologically defined components can contribute to one functionally defined component. The critical assumption is that of a consistent relation between the physiological and the functional level.

2.53 *Practical ERP component definitions*

Component definitions have generally tended to incorporate aspects of both the physiological and functional frameworks outlined above. The particular framework adopted by researchers has depended at least in part on the specific focus of the particular experimental work, and the conclusions which can legitimately be drawn from such work. Given the increasing emphasis on localisation of the generators of scalp-recorded ERPs, and the opportunities to relate ERP data with that from other neuroimaging modalities such as PET and fMRI, it seems reasonable to assume that both physiological and functional considerations will increasingly be employed to describe ERP components. In the following sections two extensively investigated endogenous ERP components - P300 and N400 - are introduced. These components are discussed briefly here since their time course and the circumstances in which they are elicited mean that they are relevant to the experimental work contained in this thesis.

2.54 *The P300 component*

The P300 is a positive-going potential, the amplitude of which is largest over central and parietal scalp locations (Sutton *et al.*, 1965). The component has a peak latency which can vary between 300 and 900 msec post-stimulus (for reviews see Coles, Gratton and Fabiani, 1990; Donchin and Coles, 1988; Pritchard, 1981). The classical task in which this component is evoked is the so-called 'odd-ball' paradigm, in which subjects respond to relatively rare target stimuli embedded in a series of non-target stimuli, to which no response is required (for example, see Donchin, Karis, Bashore, Coles and Gratton, 1986). In comparison to ERPs evoked by the non-targets, those to targets are more positive, and the size of the P300 is inversely proportional to the probability of a target (Duncan-Johnson and Donchin, 1977). The latency of the P300 has also been linked to the time taken to categorise stimuli (Coles, Gratton, Bashore, Eriksen and Donchin, 1985; Kutas, McCarthy and Donchin, 1987), whilst the amplitude of this component has been linked to the probability that a stimulus will be subsequently remembered (Karis, Fabiani and Donchin, 1984), and the confidence with which stimuli are detected (Hillyard, Squires, Bauer and Lindsay, 1971; Ruchkin and Sutton, 1978).

2.55 *The N400 component*

The N400 component was first reported by Kutas and Hillyard (1980) in a task where ERPs were recorded to the terminal words in sentences. The ERPs to terminal words which rendered the sentence meaningless were associated with a negative-going shift, which was attenuated when the terminal word fitted the context of the sentence. For example, the comparison might be between the ERPs evoked by the words *socks*, and *glass*, presented as the terminal word of the sentence '*He drank the liquor from the* _____'.

The N400 is largest over central and parietal scalp locations, and there is some evidence that in the classical sentence processing task it is larger over the right hemisphere than over the left (Kutas and Hillyard, 1982). Subsequent work has linked the amplitude of the N400 to the degree of semantic relatedness between words and their preceding context (Kutas and Hillyard, 1984), and these properties of the N400 have made it a useful tool for investigations of the on-line processing of language (for a review see Osterhout and Holcomb, 1995).

However, modulations of the N400 are not restricted to the sentence paradigm described above. For example, the N400 is attenuated by the repetition of individual words over relatively short delays (Bentin, McCarthy and Wood, 1985; Rugg, 1985; Rugg and Nagy, 1989). In addition, this component is also sensitive to non-semantic relationships between stimuli (Barrett and Rugg, 1990; Rugg and Barrett, 1987).

2.6 *Localising the neural generators of scalp-recorded ERPs*

Section 2.53 mentioned the growing emphasis on the issue of source localisation - mapping ERP components onto their neural substrate. Whilst no explicit attempt is made in the experiments reported here to localise the sources of scalp-recorded ERPs, the issue of source localisation will be discussed briefly.

At the most general level the principal obstacle for any attempt to map scalp recorded waveforms to their generators is the 'inverse' problem: for a given pattern of electrical activity recorded across the external surface of an object there is no unique solution in terms of the number and location of internal sources (Wood, 1982). One way to circumvent this problem is to assume that the brain consists of a number of discrete regions, such that a given pattern of recorded electrical data may be produced by the summated neuronal activity from within one or more of these regions (Picton *et al.*, 1995; Scherg, 1990). This assumption restricts the class of possible solutions to a finite number. The class of solutions may then be further constrained via recourse to other related data sources that speak to the relationship between neural and cognitive function (Picton, 1987). These include data from intracranial ERP recordings (Wood, 1987), neurological populations (Rugg, 1992), and other localisation techniques such as PET scanning (see Roland, Kawashima, Gulyas and O'Sullivan, 1995).

One source localisation framework which adopts some of the assumptions outlined above is the BESA technique (Brain Electrical Source Analysis; Scherg, 1990). The technique assumes that the head can be modeled as a number of spherical shells. For the description given here a model consisting of three shells will be introduced, although more recent models have proposed the use of four shells (Berg and Scherg, 1994). The three shells correspond to brain tissue, skull, and scalp. The conductivity of each layer is estimated in order to predict the spatial and temporal relationship between fields at source and at the scalp.

The BESA technique allows for the positioning and orientation of a number of equivalent dipoles, where the field generated by such a dipole is assumed to be equivalent to a field generated by an active region of brain tissue. This assumption holds if recordings are made at a sufficient distance from a neural source, and if the surface area of active tissue which a dipole approximates is small relative to the recording distance (Nunez, 1990).

One attraction of the BESA approach to localisation is that in addition to providing information concerning location and orientation the technique also estimates the temporal evolution of each dipole with respect to its contribution to the scalp-recorded waveform. The technique therefore speaks to the issue of the temporal relationship between neural activity in different brain locations. Localisation techniques such as BESA are undergoing continual refinement, and the extent to which this technique and those of a similar nature will be useful is yet to be fully realised.

2.7 *The functional significance of differences between ERPs*

The principal emphasis in the ERP analyses in the experiments reported in this thesis is on differences between ERPs evoked under different experimental conditions, and associated with different behavioural responses/outcomes. The resulting discussions focus initially on the functional implications of such differences. The issue of the neural basis for the experimental effects reported will be addressed in the general discussion sections (Chapter 11).

The following sections in this chapter discuss the necessary assumptions if functional claims are to be made on the basis of differences between ERPs evoked in any given experiment. Most of the observations follow logically from the description of the electrogenesis of ERPs and the descriptions of components given in the previous sections.

2.71 *Functional interpretations of ERP effects*

The existence of a statistically reliable difference between two ERPs indicates that some form of different neural processing was engaged in the two cases. As noted previously, in order make functional claims on the basis of such evidence it is necessary to assume a consistent relation between brain states and functional states. Put simply, the assumption is that different brain states map onto different functional states (for an extended discussion see Rugg and Coles, 1994).

However, it has also been noted that ERPs are only sensitive to an unknown proportion of the total brain activity related to a given stimulus, since activity in some cellular configurations does not propagate to the scalp. A consequence of this is that the *absence* of differences between ERPs evoked under different experimental conditions is not sufficient to support the view that functionally equivalent processes were engaged in the two cases: in either case the neural activity differentiating these hypothetical conditions may have been generated in brain tissue configured such that the activity could not be detected at the scalp, or the activity may have been too weak to be reliably detected.

An additional constraint regarding the inferences that can be made from evidence for differences in neural activity is that statements made on the basis of ERP data alone are necessarily correlational. It is always the case that the differences between two ERPs may reflect processes that are merely correlated with, or are consequential upon, those processes that are the focus of interest in a particular experiment. This is of course not a problem restricted only to ERP measures of neural activity, and the constraint applies equally to other imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), as well as intra-cerebral recording methods such as single cell recording.

A third important consideration regarding functional inferences made on the basis of differences in neural activity is the distinction between quantitative and qualitative differences in ERP activity. A quantitative difference between two ERPs refers to differences in the amplitude of some portion of the ERPs which is not accompanied by any differences in the relative distribution of the two ERPs over the scalp. Such findings are for the most part assumed to reflect the fact that the same functional process(es) are engaged in the two cases, but to differing degrees. So for example, a quantitative difference between two ERPs may reflect differences in the level of confidence associated with two task decisions.

A qualitative difference between two ERPs refers to differences in scalp distribution. This finding is commonly assumed to indicate the engagement of different functional processes - following logically from the assumption that different brain states indicate different functional states. Whilst evidence of differing scalp distributions does not demand this interpretation (Rugg and Coles, 1994), within a given experiment it is the strongest form of evidence that functionally distinct processes are in fact engaged.

The assumptions discussed above denote those which are most commonly adopted in order to make functional claims on the basis of electrophysiologically generated activity. These assumptions are implicit in the interpretations of previous ERP studies of long-term memory reviewed in chapter 3. Similarly, these same assumptions underlie the functional interpretations applied to the empirical work comprising this thesis that is reported in chapters 5-10.

Chapter 3

3 *Event-related potential studies of retrieval from memory*

Studies of the characteristics of long-term memory employing the ERP technique have largely only gained prominence over the last fifteen years. Prior to this, studies of memory in which ERPs were recorded were for the most part restricted to studies employing Sternberg's memory scanning task (Sternberg, 1966; for a review see Kutas, 1988), where subjects judge whether test items are part of a previously learned small set of target items.

Subsequent ERP studies of long-term memory can be broadly divided into studies concerned with encoding of information, and those concerned with retrieval. ERP studies of memory encoding have investigated the extent to which ERPs recorded at study predict subsequent memory performance. Differences between ERPs separated according to whether items were subsequently remembered or forgotten have been reported in a number of studies (Fabiani, Karis and Donchin, 1986; Karis *et al.*, 1984; Neville, Kutas, Chesney and Schmidt, 1986; Paller, Kutas and Mayes, 1987; Sanquist, Rohrbaugh, Syndulko and Lindsley, 1980). These differences generally take the form of more positive ERPs over the 400-800 msec time window to items that are subsequently remembered, and the effect is generally larger and more robust for recall tasks than for recognition memory tasks (Fabiani *et al.*, 1986; Karis *et al.*, 1984; Paller *et al.*, 1987; Paller, McCarthy and Wood, 1988). Whilst a considerable number of studies have reported these so-called subsequent memory or *Dm* effects (Paller *et al.*, 1987), the functional significance of the effects is unresolved.

3.1 *Indirect tests of memory*

The sensitivity of ERPs to memory-related processes has principally been investigated by the comparison of the ERPs evoked by old and new items on direct and indirect memory tests. On indirect tests, ERPs to repeated items are more positive than those to items presented for the first time, the difference between these classes of ERP onsetting 200-300 msec post-stimulus and continuing for 300-400 msec (Bentin and Peled, 1990; Rugg, Furda and Lorist, 1988; Rugg and Nagy, 1987). These effects, at least those evoked by the presentation of single words, are relatively short lived: Rugg (1990) found no evidence for ERP repetition effects when the gap between study and test exceeded 15 minutes. It has been proposed that the repetition effect for the most part reflects an attenuation of the N400 to repeated stimuli (Rugg and Doyle, 1994).

Given that the ERP repetition effect denotes changes in neural activity as a function of item repetition, it is a manifestation of some form of memory. However, the relationship between the repetition effect and the processes supporting either implicit or explicit memory is unclear. This is principally due to the fact that indirect tests provide no means of assessing subjects awareness when processing repeated and unrepeated items, and consequently no way of describing the relationship between states of awareness and the processing indexed by the ERP repetition effect.

3.2 *Direct tests of memory*

The majority of direct tests of retrieval from long-term memory in which ERPs have been recorded have employed the recognition memory paradigm. A robust finding is that in the typical recognition memory task ERPs to correctly identified old items are more positive than those to correctly identified new items (Johnson, Pfefferbaum and Kopell, 1985; Karis *et al.*, 1984; Neville *et al.*, 1986; Rugg, Brovedani and Doyle, 1992; Rugg and Doyle, 1992; Sanquist *et al.*, 1980; Smith, 1993; Smith and Guster, 1993; Smith and Halgren, 1989). Similar effects have also been reported on tests of continuous recognition memory, where items are repeated after a number of intervening items (Friedman, 1990; Friedman, 1990; Friedman and Sutton, 1987; Potter, Pickles, Roberts and Rugg, 1992; Rugg and Nagy, 1989).

Smith and Halgren (1989) identified two topographically distinct positive-going components which differentiated correctly recognised old and new items on a blocked recognition memory test. The authors interpreted the first of these components - the so called 'early' old/new effect - as an attenuation of the N400 to repeated words. Whilst the significance of the early effect is unresolved, it appears unlikely that it represents processing which differentiates remembered and forgotten items, since in a number of recognition memory studies where the period between study and test was more than a few minutes in duration there was little evidence for an early effect (Neville *et al.*, 1986; Paller and Kutas, 1992; Rugg and Doyle, 1992; Rugg and Nagy, 1989; Smith, 1993). In the study of Smith and Halgren (1989), study and test blocks consisted of only 20 items, and the interval between first and second presentation of words was relatively short.

These findings prompted the conclusion that early old/new effects are more related to study-test interval than to processes related to recognition memory (Rugg and Nagy, 1989), which is consistent with the finding that on indirect tests ERP repetition effects are absent when the period between successive presentations of the same items exceeds 15 minutes (Rugg, 1990). In the light of these considerations, the following review is restricted to 'late' old/new effects, which differentiate correctly recognised old and new words over study test intervals which can exceed 1 hour (e.g. Rugg, Cox, Doyle and Wells, 1995). Late old/new effects onset 300-400 msec post-stimulus, and from 500-800 msec are larger over the left hemisphere than over the right, in particular at temporal and parietal electrode sites (Neville *et al.*, 1986; Rugg and Doyle, 1992). Late old/new effects will hereafter be referred to simply as *old/new effects*.

Since the old/new effect indicates different processing accorded old and new test items it is a candidate for an electrophysiological index of processes related to retrieval of information from memory. However, there are a number of alternative explanations for the difference between the ERPs to old and new words. One possibility is that the ERP old/new effect is simply related to the act of making an old judgement. However, Neville *et al.*. (1986) reported that the ERPs to new words incorrectly judged old (false alarms) exhibited no old/new effect, and in the study of Rugg and Doyle (1992) the ERPs to false alarms were reliably less positive than the ERPs to words correctly judged old at latencies over which the latter ERPs exhibited an old/new effect. These findings do not support the view that the ERP old/new effect is simply a consequence of making an old decision.

A second possibility is that the ERP old/new effect reflects item repetition. If this interpretation were true, then ERPs to old items separated according to whether they were judged old or new should both exhibit an old/new effect. Bentin, Moscovitch and Heith (1992) reported such a pattern of results. However, two further studies report that ERPs to old words incorrectly judged new (misses) do not exhibit an old/new effect (Neville *et al.*, 1986; Smith, 1993). The question of differences between ERPs to misses and ERPs to words correctly judged new will be returned to in subsequent discussions.

A further factor which has been proposed to underlie the old/new effect is response confidence (Karis *et al.*, 1984): confidently detected stimuli are associated with more positive ERPs over the same time period in which old/new effects occur (Hillyard *et al.*, 1971; Ruchkin and Sutton, 1978). However, direct experimental evidence from two previous studies of recognition memory suggests that the size of the old/new effect varies little as a function of confidence (Rugg *et al.*, 1995; Rugg and Doyle, 1992). For an interpretation linking the ERP old/new effect to response confidence it is also implicitly assumed that correct old judgements to test stimuli are made more confidently than correct judgements to items presented at test for the first time. Why this should be the case is unclear (Paller and Kutas, 1992).

3.3 *Memory-related interpretations of the ERP old/new effect*

Functional interpretations relating the old/new effect to memory processes can be broadly divided into those proposing that the effect indexes processes related to retrieval from episodic memory (recollection) (Paller and Kutas, 1992; Paller, Kutas and McIsaac, 1995; Smith, 1993; Smith and Halgren, 1989; Van Petten, Kutas, Kluender, Mitchiner and McIsaac, 1991), and those which link the effect with fluency-based recognition (Friedman, 1990; Johnson *et al.*, 1985; Rugg *et al.*, 1992; Rugg and Doyle, 1994).

Smith and Halgren (1989) compared ERPs recorded during a modified recognition memory test from three groups of subjects: patients who had undergone either a right- or left- sided temporal lobectomy (R-ATL, L-ATL), and a control group of normal subjects. Subjects were exposed to six blocks of words, the first of which was a study block. For all subsequent blocks subjects made an old/new judgement to each test word. Half of the words in each test block were new words and the other half were repeated items. These repeated items were the same in each block.

Of the three groups, only the L-ATL patients did not exhibit a reliable old/new effect. However, in comparison to the controls their memory performance was only mildly impaired, and their performance as a function of repetition improved at the same rate (Smith and Halgren, 1989). The authors interpreted these findings within a dual-process model of recognition memory (Atkinson and Juola, 1973; Jacoby and Dallas, 1981). They argued that the fact that L-ATL and control subjects showed equivalent increments in performance over test blocks occurred because the L-ATL patients had an intact ability to make recognition memory judgements based on fluency. If this interpretation is correct then the absence of an old/new effect for these patients suggests that relative fluency is not indexed by ERPs. Consequently, the authors identified the old/new effect with the recollection of a prior occurrence.

The alternative view - that the old/new effect is in fact related to fluency - was proposed by Potter and colleagues (Potter *et al.*, 1992), on the basis of findings that on a continuous recognition memory task subjects displayed slightly larger old/new effects when injected with scopolamine compared to when injected with saline. Scopolamine appears to cause impairment of performance on direct tests of memory, while leaving performance on some indirect tasks unimpaired (Kopelman and Corn, 1988). Given the selective effects of scopolamine the authors assumed that whereas recollection would be impaired in the scopolamine condition, fluency-based recognition would be intact, since the processes supporting fluency have been linked to those supporting performance on indirect tests of memory (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992). Hence, compared to the control condition, in the scopolamine condition a higher proportion of correct old judgements may have been made on the basis of fluency-based recognition. The larger old/new effects in the scopolamine condition are therefore consistent with the view that the old/new effect in fact indexes fluency-based recognition.

3.31 *Recognition memory for high- and low- frequency words*

Rugg and Doyle (1992, 1994) also favoured a fluency-based interpretation of the old/new effect on the basis of findings that on tests of recognition memory low-frequency words are associated with larger old/new effects than are high- frequency words. The recognition memory advantage for low- frequency words is well documented (Gorman, 1961; Mandler *et al.*, 1982), and within dual-process theories of recognition memory it has been proposed that the advantage is due to the fact that low-frequency words are more likely to be judged old on the basis of relative fluency (Mandler, 1980; Mandler *et al.*, 1982). Rugg and Doyle (1992, 1994) employed this interpretation of the recognition memory advantage for low- frequency words to relate the old/new effect to relative fluency, arguing that the larger old/new effect for these words reflected the greater proportion of trials on which the correct old judgement had been based upon relative fluency. However, their findings are also open to an interpretation in terms of recollection, since using the R/K paradigm Gardiner and Java (1990) reported that low- frequency words attracted more R responses than did high-frequency words. The enhanced old/new effects for low- frequency words reported by Rugg and Doyle (1992, 1994) can therefore be explained by assuming that the averaged waveform for the low- frequency items was in fact composed of more responses that were accompanied by recollection of a prior encounter.

Direct support for this latter position comes from the findings of Smith (1993), who recorded ERPs to test stimuli while subjects performed an R/K recognition memory task. Qualitatively similar old/new effects were evident for both R and K responses, with the magnitude of the effect being approximately twice as large in the case of the former. These findings support the view that the old/new effect is associated with processes related to recollection. However, given the qualitatively similar old/new effects for K responses, the data are also consistent with the view that R and K responses can be described along a single dimension such as memory strength, where K responses reflect 'weak' or partial retrieval of information. The findings therefore provide no direct evidence that separate processes contribute to R and K judgements.

Paller and Kutas (1992; see also Paller, Kutas and McIsaac, 1995) linked the ERP old/new effect with recollection on the basis of the finding that, on a tachistoscopic word identification task, correctly identified words previously studied in a semantic encoding condition were associated with a larger old/new effect than words previously studied in a non-semantic encoding condition. The authors proposed that the separation of correctly identified words as a function of study task yielded response categories which differed in the proportion of words which had been recollected. They suggested that the larger ERP old/new effect for semantically encoded words reflected the fact that a higher proportion of words in this category were recollected when encountered during the identification task. The authors also argued that the data were inconsistent with an interpretation of the old/new effect in terms of fluency, since priming (the increased probability of identification of a previously studied word) was equivalent for words which had been studied either semantically or non-semantically. If the old/new effect did in fact index fluency-based recognition then equivalent old/new effects for semantically and non-semantically studied words would have been predicted.

The interpretation offered by Paller and Kutas (1992) is consistent with the findings of Rugg and colleagues (Rugg *et al.*, 1995) who investigated recognition memory for high- and low- frequency words, and for the context in which the words were encountered at study. In addition they elicited confidence judgements (high/low) to the memory decisions that were made. Using this design Rugg and colleagues were able to compare the old/new effects to high- and low- frequency words which had been correctly judged old and confidently assigned to their correct study context (the context discrimination was between words that had either been studied in a sentence generation task, or in a pleasantness rating task). Rugg and colleagues reasoned that words of this type were strong candidates for having been recollected.

The old/new effect for low- frequency words was reliably larger than the effect for high-frequency words, a finding which Rugg *et al.* (1995) interpreted in terms of graded recollection, by assuming that low- frequency words engender more recollection than do their high- frequency counterparts. Therefore, in addition to linking the old/new effect to recollection this interpretation also suggests that the old/new effect should vary systematically with either the quality or the quantity of information that is retrieved from memory.

However, in the study of Rugg *et al.* (1995) it was not possible to compare ERPs to words which were correctly judged old, and incorrectly assigned to study context. This was due to the fact that too few incorrect context judgements were made to permit formation of reliable averaged waveforms. If the ERPs evoked by this class of items were associated with equivalent old/new effects to those associated with words correctly assigned to context, the findings would be equally consistent with the view that ERPs index retrieval of information which is sufficient to support recognition judgements, but insufficient to support context judgements.

3.4 *Separating recognition with and without retrieval of context*

The studies of Smith (1993), and Rugg and colleagues (Rugg *et al.*, 1995) differ from the other studies reviewed above in that an explicit manipulation was introduced in order to identify recollected test items. In the absence of such a manipulation, the conclusions that can be drawn regarding the processing represented by the old/new effect are necessarily indirect, since there is no direct evidence indicating the basis employed for task judgements. Further, of the studies reviewed, only the study of Smith (1993) has compared ERPs to words which either were or were not recollected, and as previously noted, the results of the study provide no direct evidence for the view that more than one process contributes to recognition memory judgements.

The experiments reported in this thesis attempt to distinguish between recognition memory judgements which are accompanied or unaccompanied by recollection, employing a paradigm similar to that used by Rugg and colleagues (Rugg *et al.*, 1995). In these studies recollection is operationally defined as the ability to retrieve a salient aspect of the context of study presentation. The procedure employed in the six experiments reported consists of an initial study phase in which items are presented in one of two contexts (for instance, in experiments 3 and 4 half of the words were spoken in a male voice, and half were spoken in a female voice). In five of the six experiments[1] subjects made two judgements in a subsequent test phase - an initial old/new judgement, and for words judged old, a subsequent judgement on the context (for example, male/female voice) in which the word had been encountered at study.

[1] Experiment 5 employed a slightly different manipulation where the old/new and context judgements were both made on a single binary decision at test.

This procedure permitted the formation of two critical classes of ERPs for words correctly judged old - those associated with correct and incorrect context judgements. The rationale for this experimental procedure was that the ERPs to words correctly assigned to study context were strong candidates for having been recollected, whereas the ERPs to words associated with correct old/new judgements but incorrect context judgements may represent responses made on the basis of fluency (Jacoby and Dallas, 1981), or on the form of recognition without retrieval of context that is proposed by proponents of the declarative memory hypothesis (Squire, 1994; Moscovitch, 1994).

Comparison of the old/new effects for these two classes of ERPs to words judged old therefore offers to shed light on the processes which underlie the old/new effect. The experiments are consequently an investigation of the extent which ERPs to correct and incorrect context judgements differ in ways that are more compatible with the dual-process account of recognition memory offered by Jacoby and colleagues (Jacoby and Kelley, 1992), or the declarative memory view espoused by Squire (1982a, 1993), and Moscovitch (1992, 1994), among others.

In the review of the behavioural literature (chapter 1), three models of the relationship between recognition with and without retrieval of context were introduced: independence, redundancy, and exclusivity (see figure 1.1). The characteristics of these models can be employed to predict what patterns of ERP old/new effects would differentiate between the two views of the bases for recognition judgements which have been discussed.

The predicted pattern of ERP old/new effects if the effect in fact indexes relative fluency will be considered first. These considerations focus on the relationship between the size of any putative index of fluency in the ERPs to correctly recognised words which were subsequently either correctly or incorrectly assigned to study context. Under a model of exclusivity, any ERP index of fluency should only be evident in the ERPs to words correctly judged old and incorrectly assigned to study context. By a model of independence, any putative index of fluency should be larger in the ERPs to incorrect context judgements, since, as is evident from figure 1.1, chapter 1, the ERPs to words correctly assigned to context should contain a smaller proportion of trials which are associated with fluency-based recognition. Finally, under a model of redundancy, any ERP index of fluency should be of equal magnitude in the old/new effects evoked by correctly recognised words, irrespective of the accuracy of the subsequent context judgement. This is the case since under a redundancy model it is assumed that any recollected item would also have been judged old on the basis of the fluency associated with that item.

By all of these accounts, in comparison to the ERPs evoked by correctly recognised words which were also correctly assigned to study context, any ERP index of fluency should be at least as large in the ERPs evoked by words in which the context judgement was incorrect. However, this prediction does not hold for the declarative memory view that the processes supporting recognition are a subset of those which support retrieval of context. These processes have been characterised as those of retrieval and of integration of retrieved information (Moscovitch, 1994). Larger old/new effects for words correctly assigned to study context would be consistent with the view that ERPs index either of these putative processes, since correct recognition and retrieval of context may be associated with retrieval of more information, and/or successful integration of that information to form a representation of a prior episode.

Chapter 4

4 *General methods*

Prior to the reports of the empirical findings which address the issues raised in chapters 1 and 3, some preliminary methodological and practical considerations will be introduced. The following sections describe recording and analysis procedures common to all experiments. In the introductory sections to each experiment, procedures specific to that experiment will be noted.

4.1 *Stimulus materials*

Stimuli employed in the reported experiments consisted of words and non-words. All words were low- frequency (range 1-7 per million), open class, and drawn from the Kucera and Francis (1967) corpus. Low- frequency words were selected since they give rise to better recognition memory performance than do high- frequency words (Gorman, 1961; Mandler *et al.*, 1982). The words ranged from 4 to 9 letters in length. The pool of words from which the stimuli for each experiment were drawn is given in appendix 3.1. The pool of non-words from which stimuli were drawn is given in appendix 3.2.

4.2 *Stimulus presentation*

In each experiment a proportion of stimuli were presented auditorily, and a proportion were presented visually. Visual stimuli were presented in central vision on a TV monitor, in white letters on a black background. In the study and test phases of all experiments reported here, visual stimuli, excluding fixation cues, were displayed for a period of 300 msec. Auditory stimuli were presented through headphones at a comfortable audible level. The stimuli were stored on the hard disk of a personal computer, and they were edited so that the beginning of the stored sound segment corresponded to the onset of the spoken word.

4.3 *ERP recording*

All electrophysiological data were recorded from tin electrodes embedded in an elasticated cap. Recording locations were based on the international 10-20 system (Jasper, 1958). EOG was recorded bipolarly from additional electrodes placed on the outer canthus of the left eye, and above the supra-orbital ridge of the right eye. All channels were recorded referenced either to a single mastoid or linked mastoids (see methods sections in experimental chapters for further details).

Prior to electrode administration, for each subject the skin under the location of each electrode site was lightly abraded. This procedure reduced the electrical impedance levels at the scalp, thereby attenuating the contribution of electro-magnetic artifacts to the recorded EEG. For each subject inter-electrode impedances were always below 10 kΩ at all sites, and were below 3 kΩ in the majority of cases.

EEG was recorded on-line and each trial was stored on the hard disk of an IBM compatible PC. The data were analysed off-line after the end of the experimental session. All averaged ERPs were referenced to linked mastoids. The EOG was averaged separately for each response category to assess the influence of electro-ocular activity on the EEG data. Trials on which EOG fluctuations exceeded 120 μV were rejected prior to averaging, as were trials on which baseline drift (difference between first and last data point) exceeded 80μV. These procedures were implemented in order to eliminate trials from the experiment in which there was relatively strong evidence that a proportion of the activity in the single trial EEG was not directly related to task-specific neural activity. In order to maintain an acceptable signal/noise ratio a lower limit of 16 artifact-free trials per subject per experimental condition was also set. Subjects who contributed less than 16 trials were not included in any analyses involving that condition. This procedure is standard practice in the ERP experimental laboratory at St Andrews University.

4.4 *General analysis procedures*

4.41 *Analyses of variance*

The ERP data were principally analysed by comparing mean amplitude measurements over selected time windows. Mean amplitudes were computed relative to the amplitude of the pre-stimulus baseline (typically 100 msec in length). For all experiments initial analyses were performed on a montage of 13 scalp locations: three midline sites (Fz, Cz, Pz), left and right frontal (LF, RF, 75% of the distance from Fz to F3/F4), anterior temporal (LT, RT, 75% of the distance from Cz to T3/T4), parietal (LP, RP, 75% of the distance from Pz to P3/P4), posterior temporal (T5, T6), and occipital (O1, and O2). For ease of reference, these sites will be referred to in subsequent chapters as the *standard montage*. ERP data from other scalp sites are reported only where it adds information to that gained from the analyses over the standard montage.

For all experiments separate analyses were performed for midline and hemisphere locations. The midline analyses employed the factors of response category and electrode site, whilst for the lateral analyses the factor of hemisphere was included. Response categories are formed as a function of item type (e.g. old/new test item), and the behavioural response made to that item (e.g. old/new test judgement). All analyses of variance (for both behavioural and electrophysiological data) included the Geisser-Greenhouse correction for inhomogeneity of covariance (Keselman and Rogan, 1980, see comments below). For all analyses of ERP data, only those effects involving response category are reported, since they are the phenomena of principal interest. In each experiment, where reaction time (RT) data is reported all analyses on the RTs were also performed on the standard deviations of the reaction time distributions. The results of the analyses of the RT distributions are only reported where significant differences arise between response categories.

In all experiments, the results of the substantive analyses of variance are shown in table form. These tables show the results of all analyses involving the factor of response category. The tables display F values, p values, mean square error (MSE) values, and epsilon (ε) values. The epsilon value denotes the correction computed by the Geisser-Greenhouse formula (Keselman and Rogan, 1980). This value is multiplied by the degrees of freedom for the associated response category, thereby reducing the probability of a Type-I error. Uncorrected degrees of freedom for each factor are shown in the left-most column of each data table. Where more than one comparison is shown in the same table, uncorrected degrees of freedom are only shown for the first comparison. Where results of analyses of variance are shown in the main body of the text, the corrected degrees of freedom are given. Finally, in the tables showing the results of analyses of variance, p values greater than 0.1 are denoted by n.s., while all p values of 0.1 or less are shown, and p values of 0.05 or less are shown in bold text. All tables and ERP figures are presented at the end of the chapter in which the analyses of the data are reported.

4.42 *Analyses of onset latencies*

In addition to analyses of variance, further analyses on the ERP data were performed in some experiments in order to establish the onset of reliable differences between experimental response categories. These analyses were performed on difference waveforms obtained by subtracting the ERPs from one response category from the ERPs from a second response category. The analysis consisted of the computation of point-by-point t-statistics which were computed against the null hypothesis of no difference from baseline. The onset latency of any differences at a given channel was defined as the latency from which 10 consecutive t values were significant at the 0.05 level. Similarly, the offset of differences was defined as the latency from which 10 consecutive t values did not achieve significance at this level. This criteria was set in order to establish, with a reasonable degree of certainty, that the differences revealed were not simply due to the fact that multiple tests were performed on the data. Note that

Rugg, Doyle, and Melan (1993) employed this technique and set a criteria of 15 consecutive t-values. In their experiments the EEG was sampled every 4 msec, requiring that the minimum duration of any reliable differences revealed by the t-test analysis was 60 msec. For the experiments reported here in which t-test analyses were performed, the sampling rate was 6 msec per point. The decision to set the acceptance criterion at 10 consecutive t-values was made in order to reduce the possibility of failing to identify differences using the t-test analysis which were less than 90 msec in duration.

4.43 *Topographic analyses*

Comparisons of the scalp distribution of ERPs are also reported. These analyses were performed to establish whether the scalp distributions of ERPs associated with particular response categories were quantitatively or qualitatively different. These topographic analyses were performed across all scalp sites from which ERPs were recorded in each experiment. All topographic analyses were performed on data normalized by an algorithm proposed by McCarthy and Wood (1985). This correction is employed in order to reduce the probability of confounding differences in the size of an experimental effect with differences in the shape of the effect. Some form of correction is necessary because of the incompatibility between the assumptions of the additive model on which analyses of variance are based (Winer, 1971), and the multiplicative effect on scalp recorded EEG of changes in generator strength. This confound can be removed by computing, for each relevant condition, the size of experimental effect at each scalp site relative to the size of the effect at every other site (McCarthy and Wood, 1985).

Chapter 5

5 *An ERP study of memory for words and memory for study modality: part one*

5.1 *Introduction*

Chapter 3 introduced the general paradigm which the experiments in this thesis employ. To recap, the paradigm involves an initial study phase, and a subsequent test phase in which subjects first make an old/new judgement, and for words judged old, a second judgement on some salient contextual aspect of the prior presentation of the word. The contextual discrimination required in this initial exploratory experiment was between words that had either been seen or heard at study. In the initial study phase subjects were presented with words and non-words, half of which were presented auditorily and half of which were presented visually. In the test phase that followed subjects made initial old/new judgements and subsequent modality judgements to visually presented old and new words.

In addition to assessing the extent to which ERPs distinguish between the alternative accounts of recognition with and without retrieval of study context which were discussed in chapter 3, this experimental design permits a further, albeit indirect, means of investigating the sensitivity of ERPs to relative fluency. The fact that test presentation is visual for all items means that 50% of test items are presented in the same modality as at study, whereas 50% are presented in a different modality. Results from a number of cross-modal priming studies report that the degree of priming is greater when items are presented in the same modality at study and at test (Clarke and Morton, 1983; Jacoby and Dallas, 1981; Kirsner, Milech and Standen, 1983; Kirsner and Smith, 1974; Scarborough *et al.*, 1979). To the extent that priming is related to fluency-based recognition, then this process should be more available for recognition judgements when study and test modalities match (Jacoby, 1983; Jacoby and Dallas, 1981; Kelley, Jacoby, and Hollingshead, 1989). An interpretation linking any observed old/new effects to relative fluency would therefore be supported if the effects were larger when study and test modalities were the same (Paller and Kutas, 1992).

5.2 *Methods*

Subjects: A total of 21 subjects participated in the experiment, for which each was paid £3.00/hr. The data from 3 subjects was discarded due to excessive EOG artifact. A further two subjects were rejected because insufficient incorrect modality judgements were made to permit formation of reliable averaged waveforms for the critical response categories. Of the remaining 16 subjects, 5 were female. 13 subjects were right-handed, as defined by writing hand. Age of subjects ranged from 18 to 26 years (average age 21).

Experimental Material: Stimuli consisted of 480 words and 120 pronounceable non-words. The words were divided into four lists, each containing 120 items. Non-words were divided into two lists, each containing 60 items. Visual stimuli subtended a maximum visual angle of 1.5 degrees, and a vertical angle of 0.4 degrees. Auditory stimuli were digitised at 16 Khz with 8-bit resolution. Mean duration of auditory stimuli was 610 msec, and the stimuli were spoken by a single male voice.

Four study lists were produced. Each study list was formed by combining two of the four word lists and the two non-word lists. The word lists used were rotated across study lists such that each word list (and therefore each word) appeared on two different study lists. All non-words appeared on each study list. The study lists were subdivided into 6 blocks of 60 items. A block consisted of 40 words and 20 non-words. The order of item presentation within each block was random. Each block of 60 items was preceded by one filler item, each study list therefore consisting of a total of 366 items.

Within each block half of the items were presented auditorily and half visually. Modality of item presentation was rotated across study lists. Thus each word was presented auditorily on one list and visually on one list, while each non-word was presented auditorily on two lists and visually on two lists.

Test lists were formed by combining all four initial word lists. The resulting 480 words in each list consisted of 240 items (two initial word lists) that had been presented at study (old words), and 240 items that were presented at test for the first time (new words). Given that four study lists were employed, this procedure yielded four test lists. Of the 240 old words in each test list, half and been presented auditorily at study, and half visually.

Each test list was subdivided into 6 blocks of 80 items. A block consisted of 40 new words and 40 old words, the old words consisting of an equal number of items which had been presented either auditorily or visually at study. Two different random sequences were applied to the items within each block. This yielded a final total of eight test lists, each study list mapping onto two test lists. A filler item preceded each block of 80 test items, each test list therefore consisting of 486 items in total.

Procedure: Each subject was exposed to one of the four study lists. Following a 5 minute delay, subjects performed the test phase. At test, subjects were exposed to one of the 2 test lists that corresponded to the list they had encountered at study.

The study task was lexical decision. Following electrode placement, subjects were seated in front of the stimulus presentation monitor with the index finger of each hand resting on a microswitch. They wore a set of headphones through which auditory stimuli were presented, and were instructed that a fixation point (either an 'O' or an 'X') would serve as the cue indicating modality of presentation of the immediately following item. An 'O' indicated that the next item would be presented visually, and an 'X' indicated that the next item would be presented auditorily. Presentation of stimuli commenced 100 msec after the fixation point was removed from the screen. Subjects were instructed to respond to each item, pressing one key when a word was presented, and the other when a non-word was presented. The hands used for word and non-word decisions were counter-balanced across subjects. Subjects were informed that accuracy and speed were of equal importance. The fact that a recognition memory test would follow the study task was not mentioned. Subjects were asked to relax and to remain still during task performance. They were asked to minimise eye movements and eye blinks, with the exception of when the fixation point was present on the screen. A practice session consisting of 18 items preceded the study phase proper. The total inter-stimulus interval was 3.21 sec. Responses quicker than 400 msec, or slower than 1900 msec, were treated as errors.

At test, subjects were required to judge whether a word was old or new, and for words judged old, to report on the modality of study presentation. All 486 words were presented visually. An asterisk preceded presentation of each word, and was removed 120 msec prior to stimulus onset. Subjects were asked to make an initial old/new judgement for each word as quickly and as accurately as possible. This judgement was made on the same keys used for the lexical decision task at study. One second after this response was made a row of four question marks appeared on the screen for a duration of two seconds. For those items judged old, the question marks served as the cue for the subject to report on modality of first presentation. This decision was made on the same two keys employed for the initial old/new decision. The hands required for the first and second decision were counterbalanced across subjects such that there was no correlation between the old/new and auditory/visual judgements. Old/New responses quicker than 400 msec, or slower than 1900 msec, were treated as errors. As for the study phase, subjects were asked to restrict eye movements and eye blinks to the period when the fixation asterisk was present on the screen.

EEG Recording: EEG was recorded from the standard montage described in chapter 4. On-line sampling was at 4 msec per point for a duration of 1024 msec, commencing 120 msec prior to stimulus presentation. EOG and EEG were amplified with a bandwidth of 0.03 - 30 Hz (3db points). All channels were referred to linked mastoids.

5.3 *Results*

The following terminology will be used when referring to responses associated with test words of a particular category. Words presented visually at study, correctly identified as old, and correctly assigned to modality of study presentation will be referred to as belonging to the *visual hit/hit* response category. Words presented visually at study, correctly identified as old and incorrectly assigned to study modality will be referred to as belonging to the *visual hit/miss* response category. The corresponding terms for auditorily presented study words are *auditory hit/hit* and *auditory hit/miss*.

5.31 *Behavioural data: Study phase*

The proportions of correctly identified visually and auditorily presented words were 0.95 (s.d. = 0.05), and 0.87 (s.d. = 0.09) respectively. The proportions of correctly identified visually and auditorily presented non-words were 0.93 (s.d. = .05), and 0.80 (s.d. = 0.13) respectively. ANOVA on the behavioural data employed the factors of modality (auditory vs visual) and item type (word vs non-word). The analysis revealed main effects of both factors (respectively, $F(1,15) = 6.04$; $p < .05$, and $F(1,15) = 28.53$; $p < .001$). The main effect of modality reflected the fact that visually presented items were associated with a higher probability of correct identification than were auditorily presented items, whilst the main effect of item type reflected the fact that words were associated with a higher probability of correct identification than were non-words.

ANOVA on the RTs for study phase items also employed the factors of presentation modality and item type, and the analysis revealed main effects of both factors (respectively, $F(1,15) = 411.98$; $p < .001$, and $F(1,15) = 81.05$; $p < .001$). The main effect of modality reflected faster RTs to visually presented items (819 msec vs 1211 msec), while the main effect of item type reflected the fact that RTs for words were faster than RTs for non-words (946 msec vs 1084 msec).

5.32 *Behavioural data: Test phase*

Table 5.1 displays the probability of correct old/new response for old and new test words, separated according to study modality. For both visually and auditorily presented study words, discrimination was above chance level (visual: $t(15) = 11.37$; $p < .001$, auditory: $t(15) = 7.59$; $p < .001$). Comparison of the discrimination measures revealed an advantage for words presented visually at study ($t(15) = 3.68$; $p < .01$). These discrimination measures were formed by computing *p(hit) - p(false alarm)* (see Snodgrass and Corwin, 1988). References to measures of discrimination in subsequent analyses of behavioural results will refer to discriminations of this form unless stated otherwise.

ANOVA on the RTs for old and new words (table 5.2) employed the factors of response accuracy (correct vs incorrect) and word status (old visual vs old auditory vs new). The

ANOVA revealed main effects of both factors (respectively, $F(1.5,21.8) = 3.99$; $p < .05$, and $F(1,15) = 22.45$; $p < .001$). The main effect of response accuracy indicated faster RTs for correct responses. *Post-hoc* tests (Newman Keuls) revealed no reliable differences between the mean RTs for old visual, old auditory and new words. The largest difference between these RTs was for new words and for words presented visually at study (1443 msec and 1333 msec respectively).

Table 5.3 displays the probability of correct modality assignment for words correctly judged old, separated according to study modality. Also displayed (far right column) is the probability of assigning false alarms to the visual modality. The probability of a correct modality judgement was 0.63, a value which was above the chance level of 0.50 ($t(15) = 5.60$; $p < .001$). The probability of assigning false alarms to the visual modality was also significantly greater than 0.50 ($t(15) = 4.33$; $p < .001$).

Table 5.4 displays the RTs for initial old/new judgements, separated according to the accuracy of the subsequent modality judgement. Analyses of these RTs employed the factors of modality (auditory vs visual) and response accuracy (correct vs incorrect). The analysis revealed a significant interaction these factors ($F(1,15) = 9.98$; $p < .01$). *Post-hoc* tests (Newman Keuls) revealed that whereas the RTs for correct modality judgements did not differ as a function of study modality, incorrect modality judgements for auditorily presented study words were faster than incorrect judgements for words presented visually at study. Further, RTs to auditorily presented study words did not differ for correct and incorrect judgements, but correct modality judgements were faster than incorrect modality judgements for words presented visually at study.

5.33 *ERP Analyses*

Possible comparisons between the ERPs for the correct rejection, hit/hit, and hit/miss response categories were constrained because insufficient incorrect modality judgements were made to words presented visually at study to permit formation of reliable averaged waveforms. Two separate classes of analyses were therefore performed, the first constituting a cross-modal analysis, contrasting the ERPs associated with the correct rejection, auditory hit/hit and visual hit/hit response categories. A subsequent intra-modal analysis contrasted the ERPs associated with the correct rejection, auditory hit/hit and auditory hit/miss response categories. Analyses of the ERPs are reported for the 300-500 and 500-900 msec epochs. These time windows encompass the latency regions in which old/new effects for visually presented words have previously been reported (Paller and Kutas, 1992; Rugg and Doyle, 1992). For both the inter- and intra- modal analyses an initial global ANOVA compared the three critical response categories. Any effects involving response category were followed up by subsidiary ANOVAs comparing the critical ERPs on a pairwise basis.

Figure 5.1 displays the ERPs for the visual and auditory hit/hit response categories, and the ERPs to correct rejections. Following the initial N150 and P200 deflections the waveforms consist of two deflections, one negative with a peak latency of 500 msec, the other positive and peaking at approximately 800 msec. The waveforms for both the visual hit/hit and the auditory hit/hit response categories are more positive than those for correct rejections. This difference onsets approximately 300-400 msec post-stimulus, and is larger over the left hemisphere than the right, particularly at temporal and parietal sites. The waveforms for the two hit/hit response categories do not differ markedly from each other over the duration of the recording epoch.

5.331 *Comparison of ERPs across modality*

ANOVA of the 300-500 msec time window revealed a main effect of response category at lateral, but not at midline, electrode sites ($F(1.9,28.5) = 4.02$; $p < .05$), whilst ANOVA over the 500-900 msec window revealed main effects of response category at midline and at lateral sites (respectively, $F(1.7,25.7) = 9.78$; $p < .01$, and $F(1.6,24.6) = 11.01$; $p < .01$). The mean amplitude measures at each site for these response categories over the 300-500 and 500-900 msec epochs are shown in appendix 1.1.

The results of the subsidiary ANOVAs performed on the auditory hit/hit, visual hit/hit and correct rejection ERPs are displayed in table 5.5. Note that whilst table 5.5 displays the results of all paired comparisons of the response categories entering into the global ANOVA, the text will only refer to those analyses which revealed reliable effects involving response category in the global analyses described above.

The subsidiary ANOVAs indicated that over the 300-500 msec epoch both the visual hit/hit and auditory hit/hit ERPs were more positive than the ERPs to correct rejections at lateral sites, but were not reliably different from each other. This same pattern of results held over the 500-900 msec epoch for the analyses at midline and at lateral sites.

In the light of previous reports that the differences between the ERPs to correctly recognised old and new words are larger at left posterior than at right posterior sites (Neville *et al.*, 1986; Rugg and Doyle, 1992), two planned comparisons were performed, investigating the old/new effects for the auditory and visual hit/hit response categories at the left and right parietal sites. The separate analyses of the visual hit/hit and auditory hit/hit old/new effects revealed main effects of response category, and interactions between response category and site (visual hit/hit main effect: $F(1,15) = 24.34$; $p < .001$; visual hit/hit interaction: $F(1,15) = 5.13$; $p < .05$; auditory hit/hit main effect: $F(1,15) = 8.75$; $p < .01$: auditory hit/hit interaction: $F(1,15) = 4.89$; $p < .05$). These results reflect the fact that in both cases the old/new effects are larger at the left parietal site than at it's contralateral homologue. The mean amplitude measures at these sites can be seen in appendix 1.1.

5.332 *Comparison of ERPs within modality*

Figure 5.2 displays the ERP waveforms for the auditory hit/hit, auditory hit/miss and correct rejection response categories. The auditory hit/miss and correct rejection ERPs differ little over the duration of the recording epoch, whilst the hit/hit ERPs are more positive, the differences between these ERPs onsetting 300-400 msec post-stimulus. Mean amplitude measures for these three response categories at each electrode site are displayed in appendix 1.1. The analyses of these ERPs over both epochs revealed main effects of response category at midline and lateral sites (300-500 midline: $F(1.9,28.6) = 4.15$; $p < .05$, 300-500 lateral: $F(2,30) = 7.45$; $p < .01$, 500-900 midline: $F(1.6,24.6) = 11.38$; $p < .01$, 500-900 lateral: $F(1.5,22.7) = 10.83$; $p < .01$).

The results of the subsidiary ANOVAs performed over the 300-500 msec and 500-900 msec epochs for the auditory hit/hit, auditory hit/miss, and correct rejection ERPs are shown in table 5.6. Note that the comparison of the auditory hit/hit and correct rejection ERPs has already been described in the cross-modal analyses reported above, and the results are displayed in table 5.5. Table 5.6 shows that across both epochs the auditory hit/miss and correct rejection ERPs are not reliably different, but the auditory hit/miss ERPs are reliably less positive than the auditory hit/hit ERPs over both epochs at midline and at lateral sites.

5.333 *Analysis of misses and false alarms*

Figure 5.3 displays the ERPs for correct rejections, false alarms and misses (collapsed across modality of study presentation). Two paired ANOVAs were performed on these ERP waveforms. The first revealed no reliable differences between the ERPs to correct rejections and those to misses over the 300-500 and 500-900 msec epochs. This analysis was performed on the data from 14 subjects who made sufficient misses to permit formation of reliable averaged waveforms for this response category. The second analysis revealed no reliable differences between the ERPs to false alarms and those to correct rejections over either recording epoch. This analysis was performed over 15 subjects who made sufficient incorrect judgements to new test words[2].

5.4 *Discussion*

---

[2] In figure 5.3 the data are displayed for the 13 subjects who entered into the ERP analyses involving both the misses and the false alarms.

ERPs evoked by words correctly judged old and correctly assigned to study modality did not differ according to modality of presentation at study, but both were more positive than the ERPs evoked by correctly classified new words. The ERPs evoked by words incorrectly assigned to study modality were not reliably different from those to correctly recognised new words. These findings constrain functional interpretations of the processes indexed by ERPs on tests of recognition memory.

5.41 *The old/new effect and memory for context*

The rationale for requiring subjects correctly to assign recognised words to study modality was that a correct judgement may indicate recollection of the study episode. If this premise is accepted, then the existence of an old/new effect only for those words correctly assigned to modality is consistent with previous reports which have linked the old/new effect with the recollection of a prior occurrence (Paller and Kutas, 1992; Paller *et al.*, 1995; Smith, 1993; Smith and Halgren, 1989; Van Petten *et al.*, 1991). The data also provide little support for one alternative interpretation - that the old/new effect in fact indexes relative fluency (Friedman, 1990; Johnson *et al.*, 1985; Potter *et al.*, 1992; Rugg and Doyle, 1992).

5.42 *The old/new effect and relative fluency*

An interpretation of the data in terms of relative fluency is difficult to sustain given the absence of differences between the ERPs to incorrect modality judgements and the

ERPs to correctly classified new words. These findings suggest that the processes supporting recognition without retrieval of context are not indexed by ERPs. A second aspect of the data which does not support a link between the old/new effect and fluency-based recognition is that the auditory and visual hit/hit old/new effects were of equivalent magnitude. As previously noted, an interpretation linking the ERP old/new effect to relative fluency would have been supported if larger old/new effects were evident when study and test modalities matched.

5.43 *Old/New test judgements*

Relative to words presented auditorily at study, words presented visually were more likely to be correctly judged old at test. This finding is consistent with the transfer appropriate processing heuristic that memory performance improves with the degree to which test conditions recapitulate study context (Morris, Bransford and Franks, 1977; Roediger *et al.*, 1989; Tulving, 1983). However, within this general framework the intra-modal recognition memory advantage might still be attributable either to fluency or to recollection. To the extent that intra-modal repetitions engender more fluency than inter-modal repetitions (see previous comments on cross-modal priming), then the intra-modal advantage may reflect a relative increase in the proportion of correct old judgements made on the basis of relative fluency when study and test modalities match (Jacoby and Dallas, 1981). This interpretation is supported by the findings of Gregg and Gardiner (1994), who reported that the recognition memory for advantage for words presented in the same modality at study and at test was carried by K responses, and not by R responses.

None the less it is still possible that the intra-modal recognition memory advantage is also influenced by an increased probability of the recollection of a studied word, by virtue of the greater contextual overlap when modality is constant across study and test phases. These two accounts are of course not mutually exclusive, and the intra-modal memory advantage may reflect a relative increase in the probabilities of both recollection and fluency-based recognition when modality is the same at study and at test.

5.44 *Modality judgements*

The conditional probability of a correct modality judgement to words correctly judged old also differed according to modality of item presentation at study - visually presented study words correctly judged old were associated with a higher probability of correct modality assignment. As for old/new judgements, this advantage may in part reflect the fact that the probability of recollection - hence the probability of a correct modality judgement - is increased when modality is constant across study and test phases.

This explanation is insufficient in the present case however, unless it is assumed that subjects were guessing the study modality of all words presented auditorily at study (the probability of a correct auditory modality judgement was 0.45). This assumption is difficult to reconcile with the differences observed in the ERPs for correct and incorrect modality judgements.

One possible explanation for the intra-modal memory advantage is that it is at least in part due to the fact that when uncertain of the test modality, subjects were more likely to make a visual judgement than they were to make an auditory judgement. This proposal is supported by the fact that the probability of a visual modality judgement to a false alarm was reliably greater than chance (0.66). However, this interpretation rests upon the assumption that context judgements made to words which have been incorrectly judged old are predictive of behaviour when context judgements are made to correctly recognised old words. Batchelder and Riefer (1990) suggest that this assumption is questionable, and have proposed an alternative model for assessing response bias which is only based upon context judgements that are made to genuinely old items. This model requires subjects to discriminate between items which had previously been presented in one of three (or more) different contexts, and cannot therefore be applied to the data presented here. Further, to date no published studies have assessed whether the estimates of response bias computed from responses made to genuinely old items differ from those computed from false alarms.

5.5 *Summary*

The old/new effects observed for words correctly assigned to study modality are consistent with the view that ERPs index processes related to retrieval of contextual information. The absence of an old/new effect for words correctly judged old but incorrectly assigned to study modality suggests that ERPs do not index processes which support recognition which is unaccompanied by retrieval of contextual information. In particular, the data provide little evidence that ERPs are sensitive to processes contributing to fluency-based recognition.

Table 5.1 Probabilities of correct old/new judgements for new and old words in experiment 1. Old words are separated according to presentation modality at study (s.d. in brackets).

	Word Type		
	Visual	Auditory	New
P(Correct Judgement)	0.71(0.11)	0.64(0.14)	0.69(0.17)

Table 5.2 Reaction times (msec) for initial old/new judgements to new and old words in experiment 1. Old words are separated according to presentation modality at study.

	Response	Word Type		
		Visual	Auditory	New
RT	Correct	1283	1291	1341
SD		454	440	441
RT	Incorrect	1382	1404	1445
SD		457	443	513

Table 5.3 Probabilities of correct modality judgements for words correctly judged old in experiment 1. Also displayed (far right column) is the probability of a visual modality judgement for false alarms (s.d. in brackets).

	Word Type		
	Visual	Auditory	New
P(Correct Judgement)	0.81(0.14)	0.45(0.16)	0.66(0.15)

Table 5.4 Reaction times (msec) for initial old/new judgements in experiment 1, conditionalised on the accuracy of the subsequent modality judgement.

	Response	Study Modality	
		Visual	Auditory
RT	Correct	1249	1347
SD		433	448
RT	Incorrect	1433	1290
SD		478	426

Table 5.5 Results of pairwise analyses of the auditory (aud) hit/hit, visual (vis) hit/hit and correct rejection (CR) ERPs in experiment 1. The analyses were performed over the 300-500 msec, and 500-900 msec epochs.

	300-500 msec				500-900 msec			
Aud Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline								
Category (1,15)	6.96	2.75		**0.019**	13.28	8.84		**0.002**
Category x Site (2,30)	0.30	1.11	0.78	n.s.	2.78	1.12	0.74	0.097
Lateral								
Category (1,15)	5.73	5.66		**0.030**	11.61	15.98		**0.004**
Category x Site (4,60)	0.38	1.96	0.32	n.s.	1.37	2.25	0.38	n.s.
Category x Hem (1,15)	2.63	1.51		n.s.	1.97	7.05		n.s.
Category x Hem x Site (4,60)	0.35	0.62	0.64	n.s.	1.11	1.05	0.54	n.s.
Vis Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline								
Category	3.38	5.50		0.086	20.89	4.20		**0.001**
Category x Site	1.28	1.01	0.78	n.s.	1.57	1.69	0.72	n.s.
Lateral								
Category	6.69	8.03		**0.021**	25.87	7.02		**0.001**
Category x Site	0.24	1.23	0.34	n.s.	1.01	1.68	0.39	n.s.
Category x Hem	0.13	0.78		n.s.	3.26	3.82		0.091
Category x Hem x Site	0.54	0.27	0.63	n.s.	1.90	0.71	0.58	n.s.
Aud Hit/Hit vs Vis Hit/Hit	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline								
Category	0.00	5.55		n.s.	0.27	8.16		n.s.
Category x Site	1.78	1.32	0.79	n.s.	1.32	1.43	0.79	n.s.
Lateral								
Category	0.32	8.39		n.s.	0.00	10.34		n.s.
Category x Site	0.25	1.94	0.38	n.s.	0.14	2.34	0.43	n.s.
Category x Hem	1.27	2.20		n.s.	0.01	3.98		n.s.
Category x Hem x Site	0.64	0.55	0.73	n.s.	1.02	0.66	0.71	n.s.

Table 5.6 Results of paired analyses of the auditory (aud) hit/hit, auditory hit/miss and correct rejection (CR) ERPs in experiment 1, over the 300-500 msec, and 500-900 msec epochs. Degrees of freedom (uncorrected) as for table 5.5.

	300-500 msec				500-900 msec			
Aud Hit/Hit vs Aud Hit/Miss Midline	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	6.85	3.23		**0.019**	14.61	7.70		**0.002**
Category x Site	0.012	1.77	0.84	n.s.	0.71	1.83	0.74	n.s.
Lateral								
Category	14.57	5.53		**0.002**	13.45	17.89		**0.002**
Category x Site	0.22	2.17	0.42	n.s.	0.88	2.23	0.49	n.s.
Category x Hem	0.13	0.78		n.s.	1.38	4.20		n.s.
Category x Hem x Site	1.24	0.76	0.76	n.s.	0.88	0.98	0.77	n.s.
Aud Hit/Miss vs CR Midline	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	0.03	4.00		n.s.	0.02	3.66		n.s.
Category x Site	0.44	1.12	0.62	n.s.	2.27	1.26	0.62	n.s.
Lateral								
Category	1.99	5.43		n.s.	0.61	5.82		n.s.
Category x Site	0.45	1.14	0.48	n.s.	0.30	1.10	0.58	n.s.
Category x Hem	1.49	2.30		n.s.	0.32	5.41		n.s.
Category x Hem x Site	1.14	0.37	0.56	n.s.	0.80	0.90	0.68	n.s.

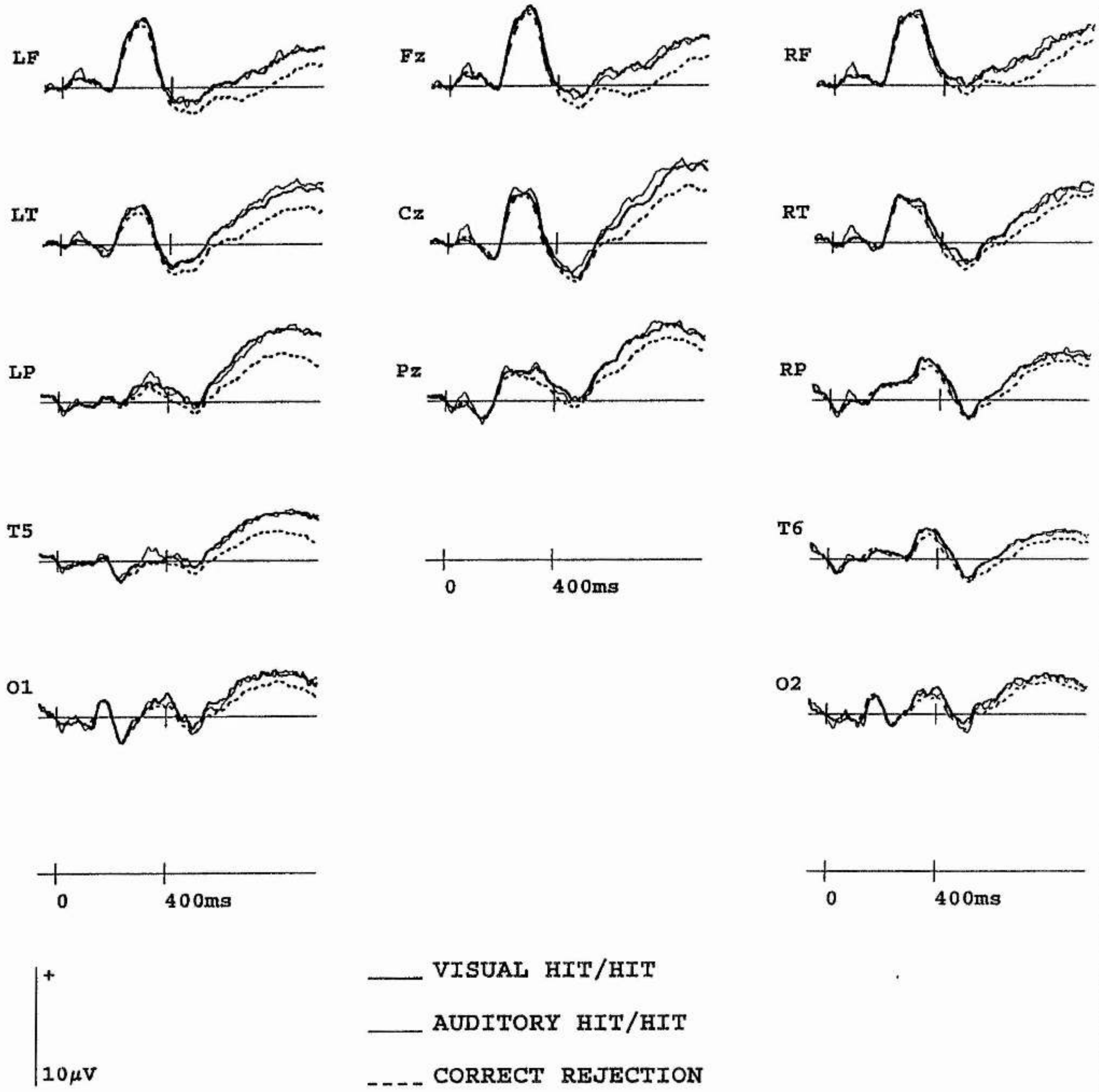


Figure 5.1 Grand average ERPs associated with the visual hit/hit, auditory hit/hit, and correct rejection response categories in experiment 1. Fz, Cz, and Pz signify midline frontal, central, and parietal sites. LF, RF, LT, RT, LP, RP, T5, T6, O1, O2 signify left and right frontal, anterior temporal, parietal, posterior temporal and occipital sites.

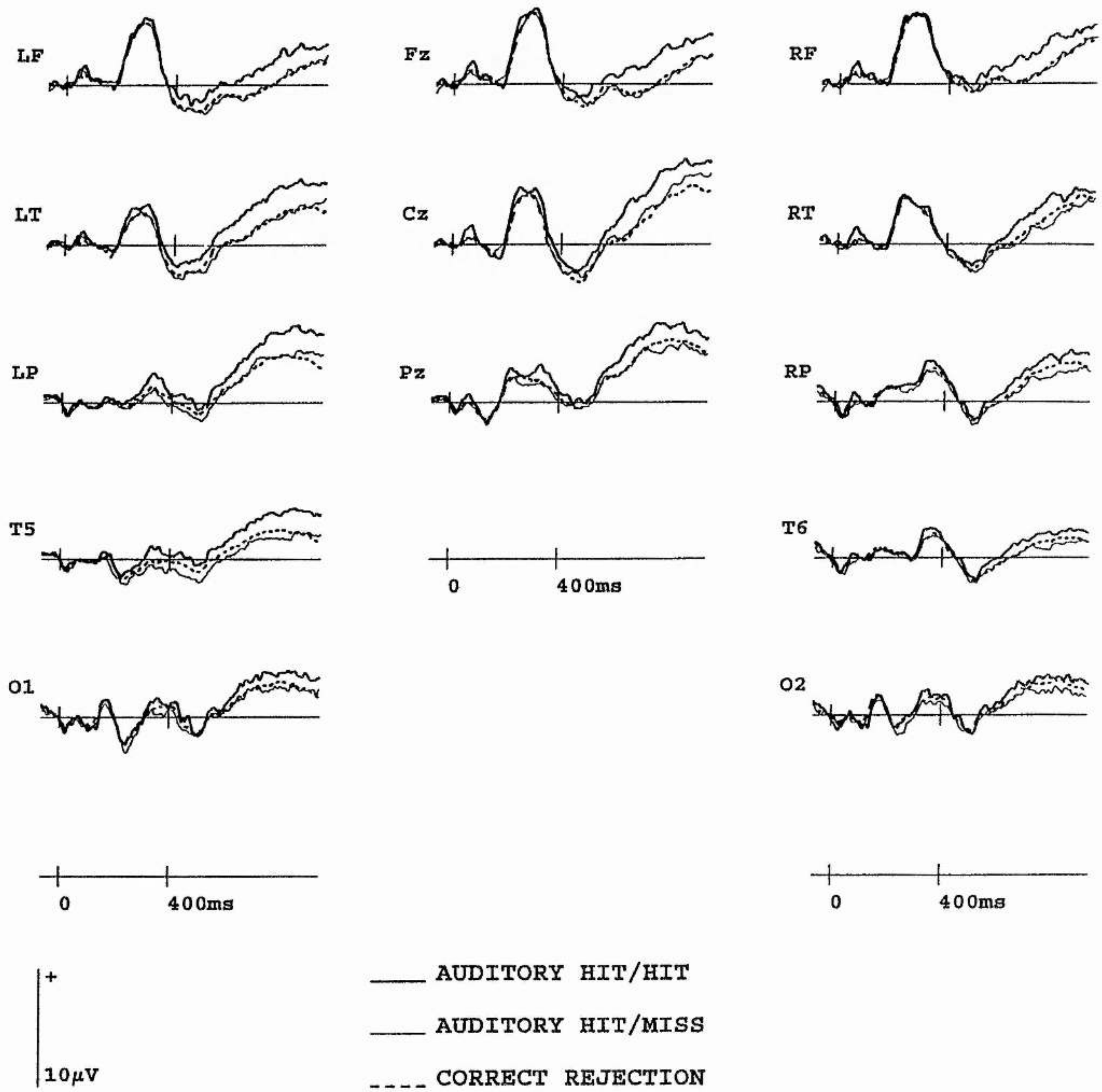


Figure 5.2 Grand average ERPs associated with the auditory hit/hit, auditory hit/miss, and correct rejection response categories in experiment 1. Electrode sites as for figure 5.1.

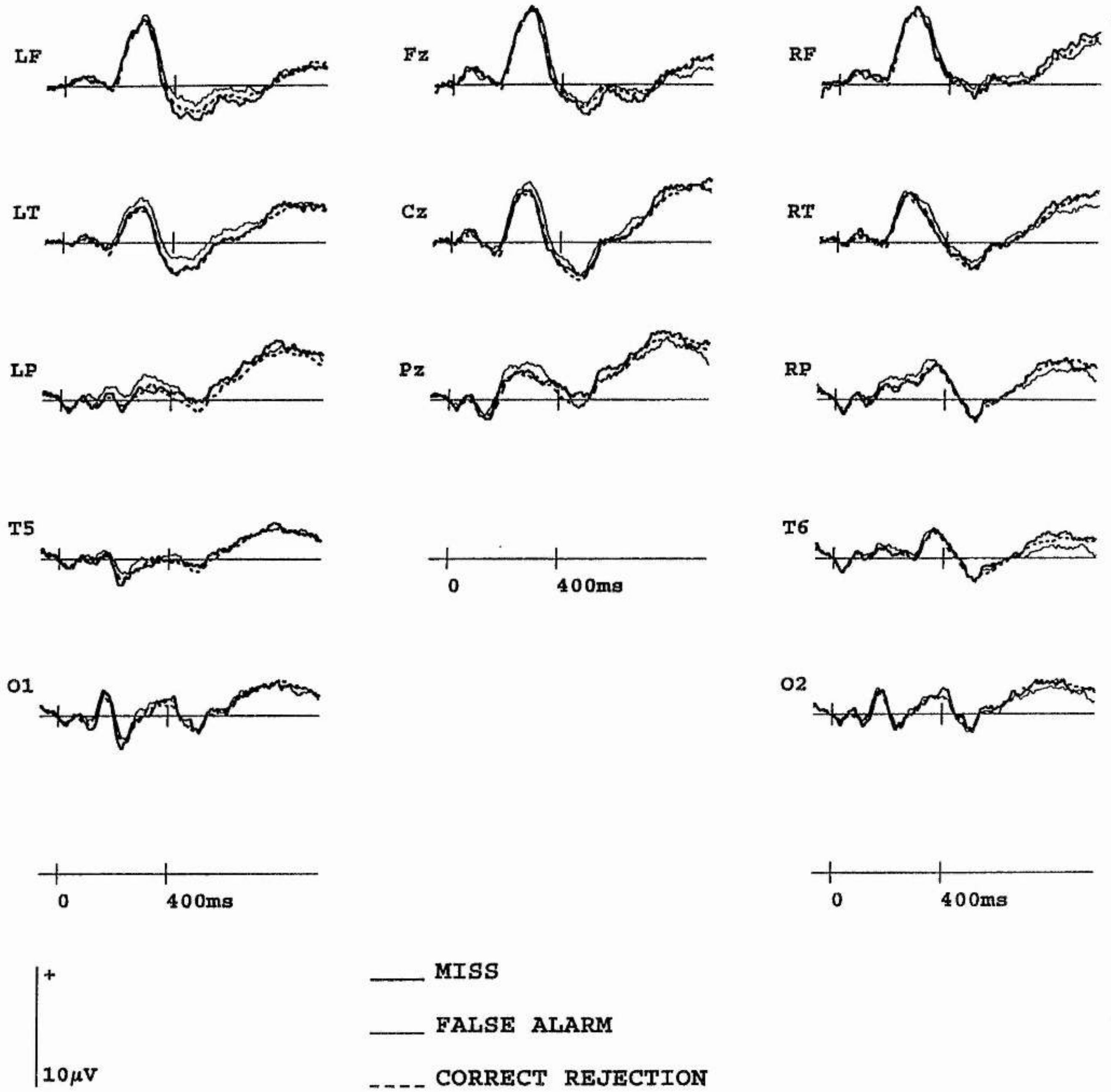


Figure 5.3 Grand average ERPs associated with the miss, false alarm, and correct rejection response categories in experiment 1. Electrode sites as for figure 5.1.

Chapter 6

6 *An ERP study of memory for words and memory for study modality: part two*

6.1 *Introduction*

The aim of experiment 2 was to generalise the findings of experiment 1, by demonstrating that ERP old/new effects were not restricted to ERPs evoked by visually presented test stimuli. Accordingly, the principal difference between this experiment and experiment 1 was that all test stimuli were presented auditorily.

6.2 *Methods*

Subjects: 19 subjects took part in the experiment, for which each was paid £3.00 /hour. For two subjects, excessive eye and head movement meant that they failed to contribute 16 artifact free trials to at least one of the critical response categories. One further subject interpreted the test instructions incorrectly. Of the remaining 16 subjects, 8 were female. 15 subjects were right handed, as defined by writing hand. Age of subjects ranged from 18 to 26 years (average age 21).

16 subjects were included in the analyses involving misses and false alarms, and the analysis of scalp distribution as a function of time. Of these 16 subjects, 1 was discarded from the analyses involving the hit/hit and hit/miss response categories due to EOG

artifact contamination. A further 2 subjects were excluded because insufficient incorrect modality judgements were made to permit formation of reliable averaged waveforms.

Experimental materials: The stimuli employed were the same as in experiment 1. The same study and test lists were employed, although the lists were randomised using different sequences. The format for the study phase of experiment 2 was identical to that for experiment 1.

At test all words were presented auditorily. An 'X' preceded each word, and was removed from the screen 100 msec before the onset of stimulus presentation. At test, responses faster than 500 msec or slower than 2200 msec were treated as errors. All other aspects of the procedure were the same as for experiment 1.

EEG Recording: EEG was recorded from the sites comprising the standard montage. On-line sampling was at 6 msec per point for a duration of 1536 msec, commencing 102 msec prior to stimulus presentation. All amplifier characteristics were as for experiment 1.

6.3 *Results*

6.31 *Behavioural data: Study phase*

At study the proportion of correctly identified visually and auditorily presented words were 0.95 (s.d. = 0.04), and 0.92 (s.d. = 0.07) respectively. The proportion of correctly identified visually and auditorily presented non-words were 0.92 (s.d. = 0.09), and 0.81 (s.d. = 0.10) respectively. ANOVA employed the factors of modality (auditory vs visual) and item type (word vs non-word). The analysis revealed main effects of both factors (respectively, $F(1,15) = 15.42$; $p < .01$, and $F(1,15) = 15.12$; $p < .01$), and in addition the analysis revealed a significant interaction term ($F(1,15) = 8.44$; $p < .05$). *Post-hoc* tests (Newman Keuls) revealed that words were associated with a higher probability of correct identification than non-words for auditorily presented items, but not for items presented visually. Further, visually presented items were associated with a higher probability of correct identification for non-words, but not for words.

ANOVA of the study phase RTs also employed the factors of presentation modality and item type. The analysis revealed main effects of both factors (modality: $F(1,15) = 800.36$; $p < .001$, item type: $F(1,15) = 118.61$; $p < .001$). The main effect of modality reflected faster RTs to visually presented items (805 msec vs 1173 msec). The main effect of item type reflected the fact that words were associated with faster RTs than were non-words (918 msec vs 1059 msec).

6.32 *Behavioural Data: Test Phase*

Table 6.1 displays the probability of correct response for old and new test words, separated according to study modality. Discrimination was above chance level for both auditorily and visually presented study words (for visual hits, $t(15) = 11.40$; $p < .001$, for auditory hits, $t(15) = 14.45$; $p < .001$). Comparison of the discrimination measures revealed an advantage for words presented auditorily at study ($t(15) = 3.98$; $p < .01$).

ANOVA on the RTs for old and new items (table 6.2) employed the factors of response accuracy (correct vs incorrect) and item status (old visual vs old auditory vs new). The ANOVA revealed main effects of both factors (respectively, $F(1,15) = 25.40$; $p < .001$, and $F(1.6,24.5) = 12.15$; $p < .001$). The effect of response accuracy reflected faster RTs for correct responses. *Post-hoc* tests (Newman Keuls) revealed that whereas the mean RTs for visually and auditorily presented study words did not differ, both were associated with significantly faster RTs than were new words.

Table 6.3 displays the probability of a correct modality judgement, separated according to presentation modality at study. Also displayed (far right column) is the probability of assigning false alarms to the visual modality. The overall probability of a correct modality judgement was 0.76, which was significantly greater than the chance level of 0.50 ($t(15) = 9.45$; $p < .001$). The probability of assigning false alarms to the visual modality was not reliably different from chance.

ANOVA of the RTs for the hit/hit and hit/miss response categories (table 6.4) employed the factors of modality of study presentation (auditory vs visual) and response accuracy (correct vs incorrect). The ANOVA revealed main effects of both factors (respectively, $F(1,15) = 7.76$; $p < .05$, and $F(1,15) = 16.03$; $p < .01$). The main effect of modality reflected the fact that words presented auditorily at study were associated with faster RTs at test than were words presented visually. The main effect of response accuracy reflected the fact that correct modality judgements were associated with faster RTs than were incorrect modality judgements.

6.33 *ERP Analyses*

As in experiment 1, comparisons involving the hit/hit and hit/miss response categories were constrained by the fact that, for both auditorily and visually presented study words, insufficient incorrect modality judgements were made to permit formation of reliable averaged waveforms. Consequently, three separate analyses involving the hit/hit and hit/miss ERP waveforms were performed. The first analysis was cross-modal, comparing the auditory and visual hit/hit ERPs with the ERPs associated with correct rejections. Two further analyses were performed in which the hit/miss ERPs were collapsed across modality of initial presentation by computing weighted averages of the ERPs associated with auditory and visual study presentation. In the first of these analyses, the ERPs for the collapsed hit/miss response category were compared to the ERPs for the auditory hit/hit and visual hit/hit response categories. In the second analysis, the ERPs associated with the collapsed hit/miss response category were compared to the ERPs associated with correct rejections. As for experiment 1, where three or more response categories were compared, any effects involving response category were followed up by subsidiary pairwise ANOVAs.

Figure 6.1 displays the ERP old/new effects for the auditory and visual hit/hit response categories. Following the initial N150 and P200 deflections, the ERPs consist of two principal deflections, one negative with a peak latency of approximately 500-600 msec, the other positive and peaking at approximately 1000-1100 msec. The ERPs for the hit/hit response categories are more positive than the ERPs to correct rejections, this difference onsetting approximately 400 msec post-stimulus (later at posterior sites). From approximately 1100 msec post-stimulus, the ERPs for the visual hit/hit response category continue to be more positive than the ERPs to correct rejections, whilst the differences between the auditory hit/hit and correct rejection ERPs diminish.

Figure 6.2 displays the ERPs for the hit/hit, hit/miss and correct rejection response categories. For clarity the hit/hit response categories have been collapsed across study modality. The hit/hit and the hit/miss ERPs are more positive than the ERPs to correct rejections between 400 and 800 msec post-stimulus. From approximately 800-1300 msec post-stimulus the hit/hit ERPs continue to be more positive than the ERPs to correct rejections, whereas the hit/miss and correct rejection ERPs cease to differ.

The ERPs displayed in figures 6.1 and 6.2 were analysed over 3 latency regions: 400-800, 800-1100, and 1100-1400 msec. These time windows were principally selected on the basis of visual inspection of the ERPs, given the lack of published data for ERP old/new effects to auditorily presented test stimuli (but see Domalski, Smith and Halgren, 1991). Mean amplitude measures for the ERPs to the correct rejection, hit/hit, and collapsed hit/miss response categories are displayed in appendix 1.2.

6.331 *Comparison of ERPs across modality*

Analysis of the 400-800 msec and 800-1100 msec time windows revealed main effects of response category at both midline and lateral sites (400-800 midline: $F(1.9,22.6) = 11.56$; $p < .001$, 400-800 lateral: $F(1.6,18.8) = 7.51$; $p < .05$, 800-1100 midline:

$F(1.8,21.4) = 8.85$; $p < .01$, 800-1100 lateral: $F(1.7,21.0) = 11.43$; $p < .001$). The results of the subsidiary ANOVAs investigating the effects revealed in the global analyses are displayed in table 6.5. These subsidiary analyses revealed that whilst the auditory and visual hit/hit ERPs did not differ over the 400-800 and 800-1100 msec epochs, both were more positive than the ERPs to correct rejections at midline and at lateral sites.

The global analysis over the 1100-1400 msec epoch revealed a main effect of response category at lateral sites ($F(1.9,22.3) = 7.34$; $p < .01$). The subsidiary ANOVAs revealed that whilst the correct rejection and auditory hit/hit ERPs did not differ, both were more negative than the ERPs for the visual hit/hit response category.

6.332 *Comparison of ERPs to correct and incorrect modality judgements*

The auditory and visual hit/hit ERPs are displayed in figure 6.1, whereas the hit/miss ERPs are displayed in figure 6.2. Analysis of the 400-800 msec time window revealed no differences between these ERPs at either midline or lateral sites. The initial ANOVA comparing these ERPs over the 800-1100 msec time window revealed a main effect of response category at midline ($F(1.7,20.8) = 4.66$; $p < .05$), and at lateral sites ($F(1.9,22.6) = 3.93$; $p < .05$). The subsidiary ANOVAs (table 6.6) revealed that at lateral sites both the auditory hit/hit and visual hit/hit ERPs were more positive than the hit/miss ERPs. At the midline, the auditory hit/hit ERPs are more positive than the hit/miss ERPs, whilst the differences between the visual hit/hit and hit/miss ERPs, although in the same direction, were not significant.

ANOVA over the 1100-1400 msec time window once more revealed a main effect of response category at midline and lateral sites (midline: $F(1.8,22.0) = 4.32$; $p < .05$; lateral: $F(1.9,22.3) = 8.65$; $P < .01$). Subsidiary ANOVAs revealed that the visual hit/hit ERPs are more positive than the hit/miss ERPs at midline and lateral sites, whereas the auditory hit/hit and collapsed hit/miss ERPs are not reliably different.

The results of the comparison of the hit/miss and correct rejection ERPs are displayed at the bottom of table 6.6. The tabulated values indicate that for both midline and lateral sites the hit/miss ERPs are more positive over the 400-800 msec epoch, but that these ERPs are not reliably different over the later epochs.

6.333 *Analysis of misses and false alarms*

As for experiment 1, ERPs to false alarms and to misses (collapsed across study modality) were compared to the ERPs for correct rejections. These ERPs are displayed in figure 6.3. In separate analyses the ERPs to misses and the ERPs to false alarms were compared to those for correct rejections. These analyses were performed over the same time windows employed for the analysis of hit/hit and hit/miss old/new effects reported above, and revealed no reliable differences at either midline or lateral scalp locations.

6.334 *Analysis of scalp distribution*

Two analyses of scalp distribution were performed on the data from experiment 2. For these analyses the hit/hit ERPs were collapsed across study modality in order to improve the signal/noise ratio. The first analysis of scalp distribution compared the collapsed hit/hit and hit/miss ERPs between 400 and 800 msec - the epoch over which reliable old/new effects were evident for the ERPs to both response categories. The analysis revealed no interaction between response category and site, which is consistent with the view that the neural generators underlying these effects are equivalent. The second analysis compared the scalp distribution of the differences between the hit/hit and correct rejection ERPs across the 400-800 and 800-1100 msec epochs. This analysis was performed on the difference waveforms obtained by subtracting the correct rejection from the hit/hit ERPs. The analysis revealed no interaction between epoch and site, which is again consistent with the view that the same generators contribute to the old/new effects across these successive epochs.

6.4 *Discussion*

Consistent with the findings of experiment 1, ERPs evoked by words correctly assigned to study modality were more positive than ERPs evoked by correctly classified new words. In contrast to experiment 1, the ERPs evoked by words incorrectly assigned to study modality were also more positive than those to words correctly judged new. This old/new effect was restricted to the 400-800 msec latency region, whereas the old/new effects for words correctly assigned to study modality were more extended in time.

6.41 *The old/new effect and relative fluency*

The differences between the ERPs to old and new words from 400-800 msec are candidates for an electrophysiological index of processes related to recognition which is unaccompanied by retrieval of study context, since over this epoch both the hit/hit and hit/miss ERPs were associated with old/new effects. However, there is only qualified support for the view that these effects reflect processes related to fluency-based recognition. Since the ERP old/new effects are of equivalent magnitude for the hit/hit and hit/miss ERPs they are consistent with an interpretation in terms of fluency only if a relationship of redundancy holds between fluency and recollection (see chapter 3).

There are also two lines of evidence which suggest that the old/new effects observed over this epoch do not reflect processes related to fluency. First, as in experiment 1, an interpretation linking these effects to relative fluency would have been supported if the old/new effects were larger when study and test modalities matched. The comparison of the auditory and visual hit/hit ERPs revealed no reliable differences over the 400-800 msec epoch. Second, the analyses of scalp topography suggested that the same neural/functional processes contributed to the ERP old/new effects across the 400-800 and 800-1100 msec epochs. Given that over the latter epoch old/new effects were evident only for words correctly assigned to study modality, the data provide little support for the view that these old/new effects reflect processes related to fluency-based recognition.

6.42 *Graded retrieval of information*

The experimental findings are consistent with the view that recognition with and without retrieval of contextual information depends upon a common process or processes, and that successful and unsuccessful retrieval of context is differentiated by the duration of this process. The data are therefore consistent with the view that a subset of those processes which contribute to retrieval of context also contribute to recognition (Squire and Zola-Morgan, 1988).

However, the findings are also consistent with the view that recognition with and without retrieval of context can be described along a single dimension of memory strength. By this view the temporally restricted old/new effect for the hit/miss ERPs denotes retrieval of information sufficient to make a correct old/new recognition judgement, but insufficient to make a correct context judgement. The temporally extended old/new effect for the hit/hit ERPs denotes retrieval of information sufficient to make a correct old judgement, and to place a test word in its correct context.

This interpretation relates graded retrieval to the duration of old/new effects. Alternatively, graded retrieval may be reflected in the amplitude of old/new effects. An amplitude based interpretation of graded old/new effects has recently been advanced by Rugg *et al.* (1995). It is not possible to separate these rival interpretations on the basis of the current data.

6.43 *Old/new effects for the hit/hit ERPs*

The analysis of the old/new effects for the auditory and visual hit/hit ERPs revealed that the effect for words presented visually at study was more temporally extended. If ERPs are in fact sensitive to processes related to recollection, as has been suggested, then it is not unreasonable to suppose that ERPs will differentiate retrieval of different forms of information, such as whether a test word was seen or heard. It is unclear however why the old/new effect for words presented visually at study and correctly assigned to modality at test should be associated with an old/new effect which is temporally extended compared to the effect associated with words presented auditorily at study.

6.44 *Behavioural data*

Relative to words presented visually at study, words presented auditorily were more likely to be correctly judged old at test, replicating the intra-modal recognition memory advantage found in experiment 1. In contrast to the findings of the previous experiment, the probability of a correct context judgement did not vary according to modality of study presentation. In this experiment there was also no evidence for a response bias for the context judgement: the probability of a visual modality judgement to a false alarm was 0.45. These findings are consistent with the view that the increased probability of a correct context judgement in experiment 1 to words presented in the same modality at study and at test was in part due to a bias towards responding 'visual' when uncertain of study context.

6.5 *Summary*

Whilst varying in duration, the ERP old/new effects to correctly recognised words which attracted either a correct or an incorrect modality judgement were morphologically similar, suggesting that the same process(es) contribute to recognition with and without retrieval of context. The findings of this experiment, and of experiment 1, are consistent with the view that these two forms of memory are differentiated by either the quality or quantity of information retrieved from memory. The findings of both experiments provide little evidence supporting the view that ERPs are sensitive to processes related to relative fluency.

Table 6.1 Probabilities of correct old/new judgements for new and old words in experiment 2. Old words are separated according to presentation modality at study (s.d. in brackets).

	Word Type		
	Visual	Auditory	New
P(Correct Judgement)	0.68(0.13)	0.75(0.11)	0.79(0.09)

Table 6.2 Reaction times (msec) for initial old/new judgements to new and old words in experiment 2. Old words are separated according to presentation modality at study.

	Response	Word Type		
		Visual	Auditory	New
RT	Correct	1509	1468	1567
SD		392	392	436
RT	Incorrect	1650	1660	1748
SD		445	449	464

Table 6.3 Probabilities of correct modality judgements for words judged old in experiment 2. Also displayed (far right column) is the probability of a visual modality judgement for false alarms (s.d. in brackets).

	Word Type		
	Visual	Auditory	New
P(Correct Judgement)	0.76(0.10)	0.75(0.15)	0.45(0.17)

Table 6.4 Reaction times (msec) for correct old judgements in experiment 2, separated according to the accuracy of the subsequent modality judgement.

	Response	Study Modality	
		Visual	Auditory
RT	Correct	1478	1450
SD		368	383
RT	Incorrect	1588	1507
SD		394	396

Table 6.5 Results of pairwise analyses of the auditory (aud) hit/hit, visual (vis) hit/hit and correct rejection (CR) ERPs in experiment 2. The analyses were performed over the 400-800, 800-1100, and 1100-1400 msec epochs.

	400-800 msec				800-1100 msec				1100-1400 msec			
Aud Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category (1,12)	20.72	3.21		**0.001**	21.58	3.91		**0.001**	0.92	5.72		n.s.
Category x Site (2,24)	0.29	0.59	0.68	n.s.	0.46	1.74	0.69	n.s.	2.98	2.25	0.64	0.10
Lateral												
Category (1,12)	8.69	7.97		**0.012**	29.83	5.47		**0.001**	0.06	13.32		n.s.
Category x Site (4,48)	0.44	1.27	0.37	n.s.	0.94	2.01	0.36	n.s.	0.61	2.33	0.41	n.s.
Category x Hem (1,12)	0.16	1.68		n.s.	0.20	1.47		n.s.	0.97	2.55		n.s.
Category x Hem x Site (4,48)	1.36	0.38	0.60	n.s.	0.91	0.89	0.52	n.s.	1.05	1.56	0.44	n.s.
Vis Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	14.33	2.32		**0.003**	5.75	6.67		**0.003**	0.77	9.33		n.s.
Category x Site	0.16	0.64	0.78	n.s.	0.40	0.85	0.97	n.s.	0.03	1.66	0.80	n.s.
Lateral												
Category	16.90	3.06		**0.001**	12.27	10.91		**0.004**	9.53	16.43		**0.040**
Category x Site	3.24	0.81	0.54	0.52	0.48	1.40	0.46	n.s.	0.33	1.93	0.55	n.s.
Category x Hem	0.07	2.13		n.s.	0.01	2.69		n.s.	0.00	3.76		n.s.
Category x Hem x Site	0.21	0.34	0.53	n.s.	0.78	0.63	0.58	n.s.	0.81	1.20	0.51	n.s.
Aud Hit/Hit vs Vis Hit/Hit	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	1.60	3.60		n.s.	2.08	4.29		n.s.	4.47	5.53		0.056
Category x Site	0.37	1.30	0.58	n.s.	0.54	1.31	0.72	n.s.	4.17	1.83	0.71	**0.046**
Lateral												
Category	0.25	5.25		n.s.	0.15	9.72		n.s.	13.39	10.09		**0.003**
Category x Site	1.01	2.03	0.45	n.s.	0.30	2.41	0.59	n.s.	0.72	3.62	0.61	n.s.
Category x Hem	0.51	2.18		n.s.	0.04	3.69		n.s.	0.46	4.89		n.s.
Category x Hem x Site	0.78	0.58	0.65	n.s.	0.85	0.93	0.64	n.s.	0.69	1.21	0.64	n.s.

Table 6.6 Results of paired comparisons in experiment 2 of the hit/miss ERPs with the Auditory (aud) hit/hit, Visual (vis) hit/hit, and correct rejection (CR) ERPs respectively. Analyses were performed over the 400-800 msec, 800-1100, and 1100-1400 msec epochs. Degrees of freedom (uncorrected) are as for table 6.5.

	400-800 msec				800-1100 msec				1100-1400 msec			
Aud Hit/Hit vs Hit/Miss Midline	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	0.00	9.93		n.s.	8.69	7.11		**0.012**	0.96	5.18		n.s.
Category x Site	0.61	0.81	0.85	n.s.	0.03	2.23	0.72	n.s.	0.85	1.69	0.71	n.s.
Lateral												
Category	0.08	24.01		n.s.	5.42	14.90		**0.038**	0.08	14.64		n.s.
Category x Site	1.13	2.05	0.36	n.s.	0.44	3.38	0.36	n.s.	0.36	3.03	0.42	n.s.
Category x Hem	1.40	1.57		n.s.	0.05	2.99		n.s.	0.01	2.31		n.s.
Category x Hem x Site	0.33	0.60	0.74	n.s.	0.09	0.90	0.76	n.s.	0.14	1.43	0.71	n.s.
Vis Hit/Hit vs Hit/Miss Midline	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	0.40	13.94		n.s.	2.68	8.86		n.s.	6.35	8.17		**0.027**
Category x Site	0.50	1.20	0.65	n.s.	0.43	1.58	0.70	n.s.	0.90	2.81	0.58	n.s.
Lateral												
Category	0.29	22.13		n.s.	5.17	11.70		**0.042**	16.70	9.66		**0.002**
Category x Site	0.61	1.94	0.46	n.s.	0.40	3.10	0.43	n.s.	0.27	3.15	0.46	n.s.
Category x Hem	0.07	2.51		n.s.	0.15	3.94		n.s.	0.49	5.57		n.s.
Category x Hem x Site	0.78	0.79	0.68	n.s.	1.09	1.21	0.62	n.s.	0.12	1.90	0.63	n.s.
Hit/Miss vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	6.65	9.92		**0.024**	0.19	9.16		n.s.	2.30	8.92		n.s.
Category x Site	0.52	1.17	0.65	n.s.	0.54	2.02	0.61	n.s.	0.97	2.02	0.73	n.s.
Lateral												
Category	4.81	19.58		**0.049**	0.71	20.24		n.s.	0.00	18.39		n.s.
Category x Site	1.44	1.18	0.42	n.s.	1.25	1.34	0.54	n.s.	0.15	1.14	0.68	n.s.
Category x Hem	0.01	0.95		n.s.	0.44	2.01		n.s.	0.78	3.78		n.s.
Category x Hem x Site	1.32	0.40	0.74	n.s.	1.65	0.77	0.63	n.s.	0.92	1.11	0.65	n.s.

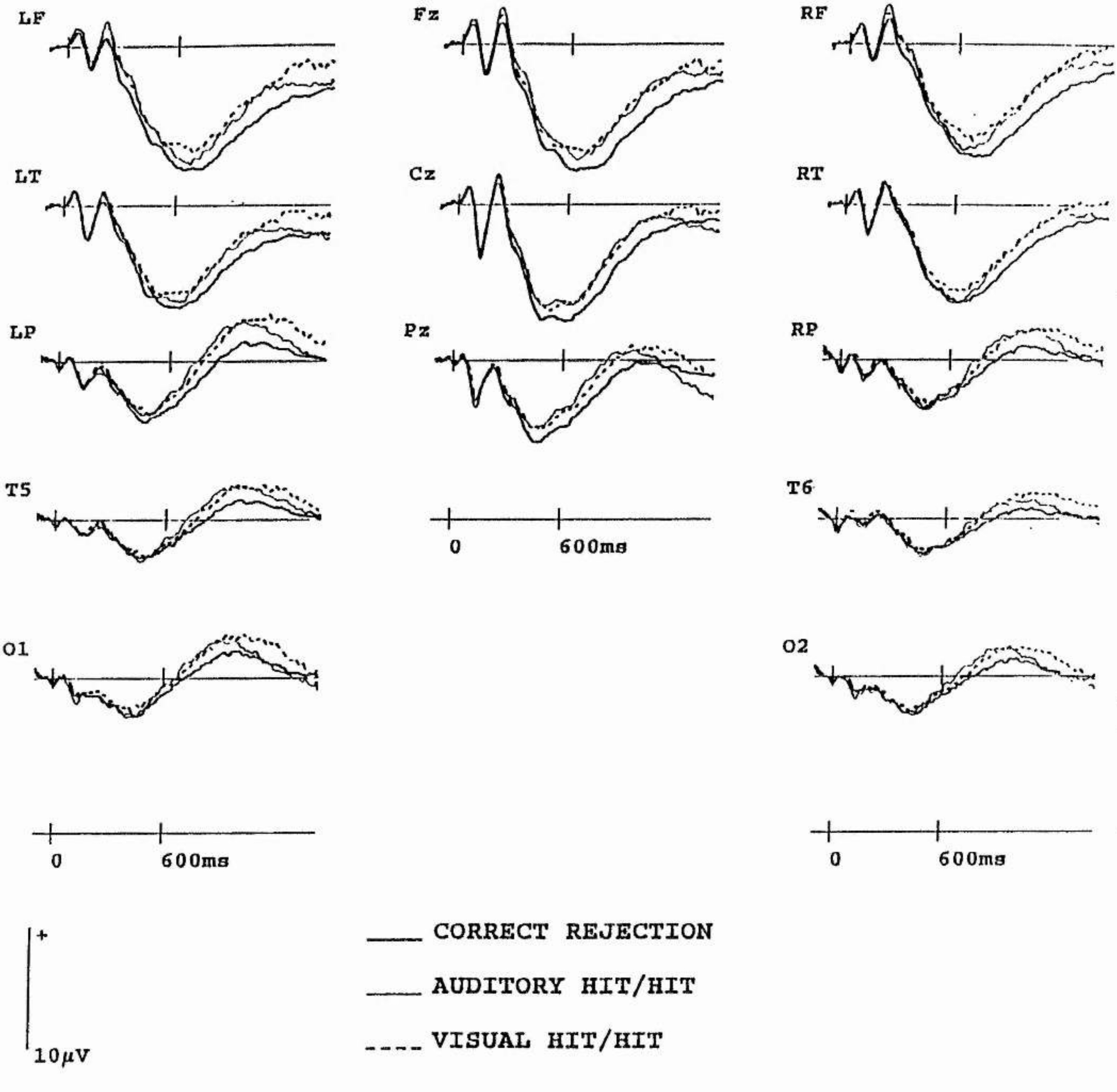


Figure 6.1 Grand average ERPs associated with the visual hit/hit, auditory hit/hit, and correct rejection response categories in experiment 2. Electrode sites as for figure 5.1.

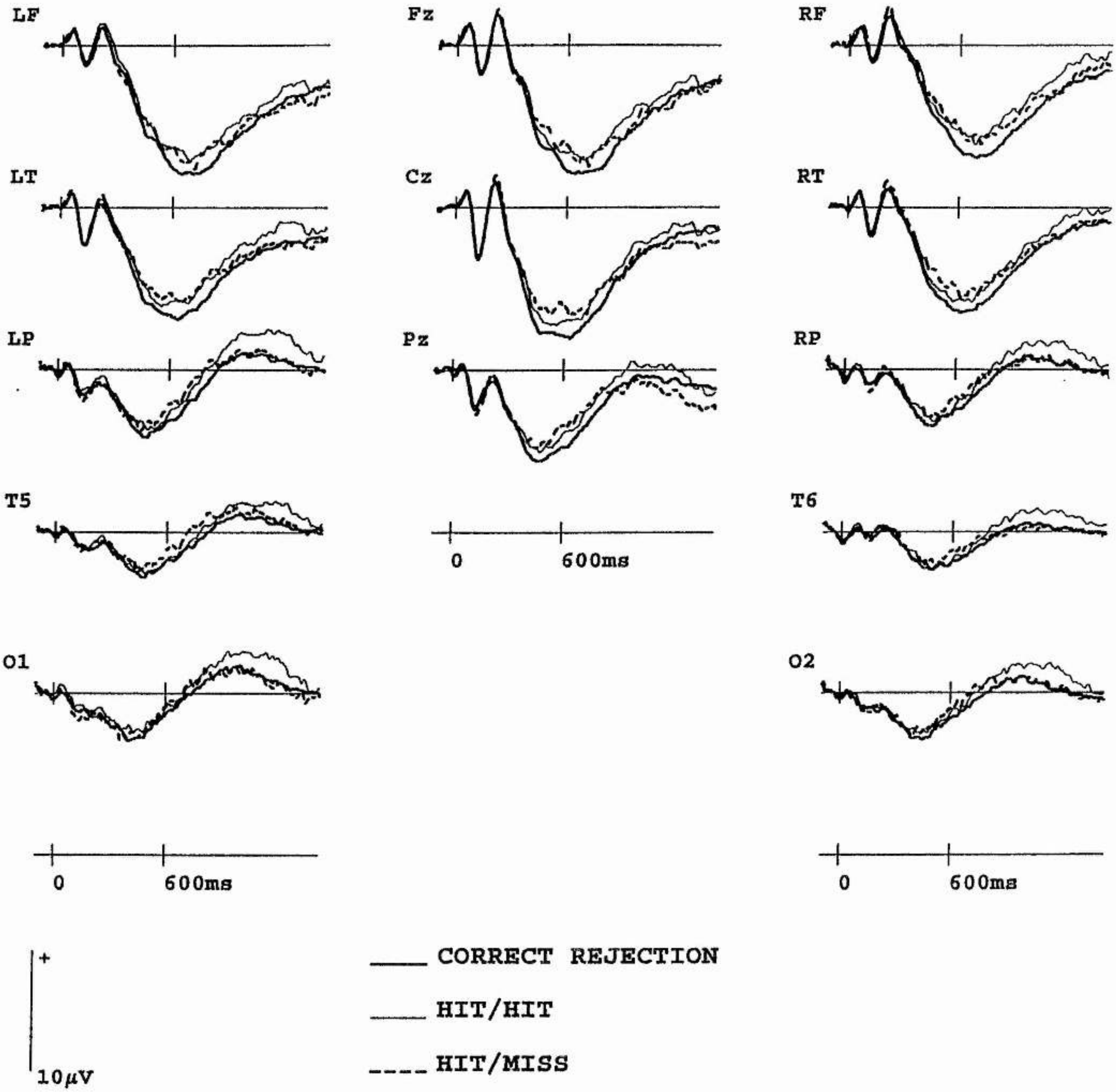


Figure 6.2 Grand average ERPs associated with the hit/hit, hit/miss, and correct rejection response categories in experiment 2. Electrode sites as for figure 5.1.

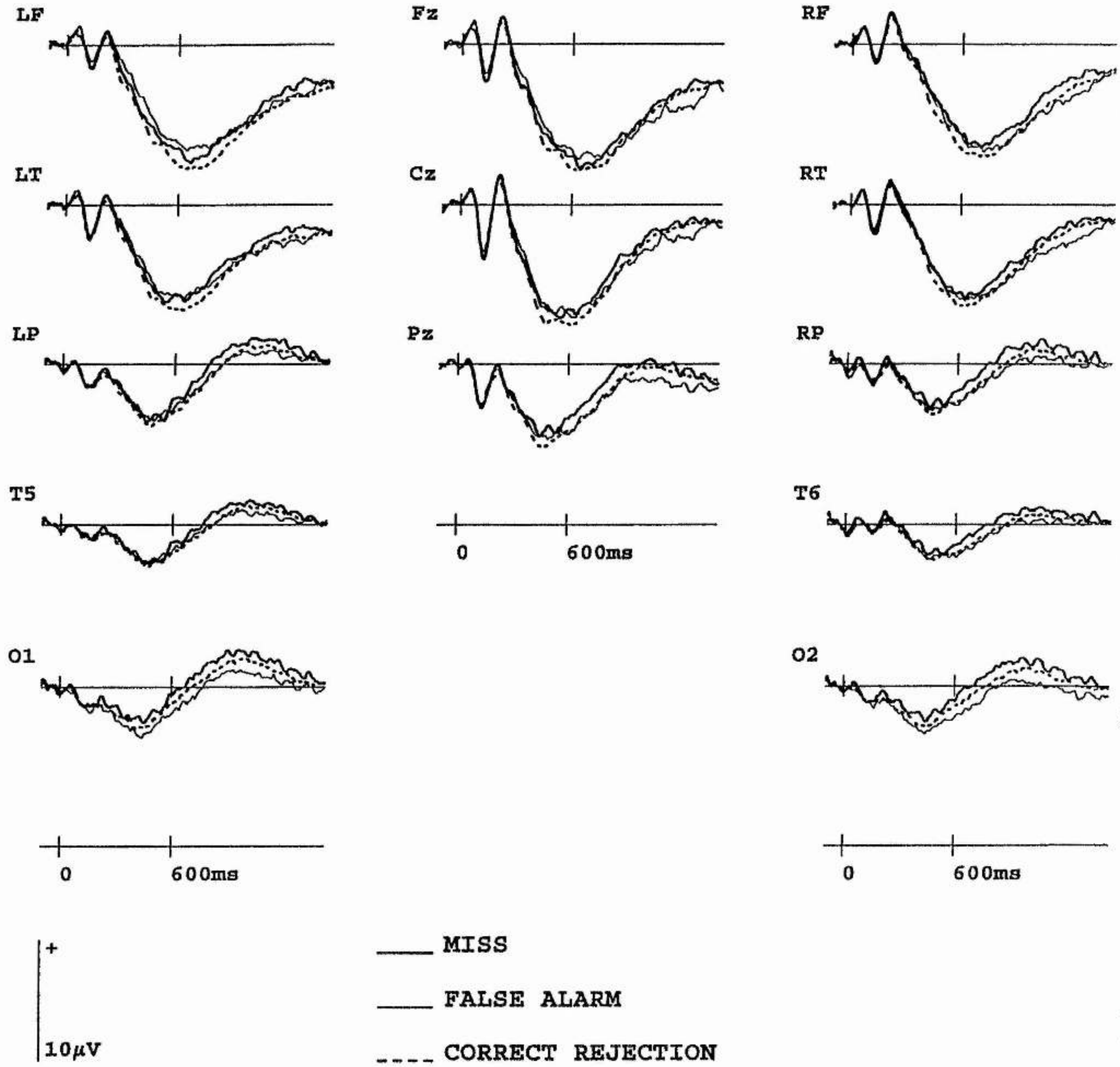


Figure 6.3 Grand average ERPs associated with the miss, false alarm, and correct rejection response categories in experiment 2. Electrode sites as for figure 5.1.

Chapter 7

7 *An ERP study of memory for words and memory for speaker voice: part one*

7.1 *Introduction*

The interpretations offered for the results from experiments 1 and 2 are tempered by a potential confound introduced by the use of modality as the marker of retrieval of study context. It has been noted previously that in both of these experiments 50% of old test words were presented in the same modality as at study, whereas 50% were presented in a different modality. Whilst this facet of the design was employed to argue that any ERP index of fluency should be larger when study and test modalities were the same, it has also been suggested that, in the absence of recollection, the fluency with which test items are processed can serve as a basis for modality judgements (Kelley *et al.*, 1989).

If this suggestion is correct, then the ERPs to putatively 'recollected' words in these two experiments may have included an unknown proportion of words on which the modality judgement was based upon fluency rather than recollection. The differences between the ERP old/new effects for correctly recognised words which were correctly or incorrectly assigned to study modality may therefore not be an accurate representation of any different neural processing associated with recognition which is accompanied or unaccompanied by retrieval of study context.

7.11 *Latency jitter*

A second consideration relevant to the findings of experiment 2 is that auditory presentation at test may have reduced the likelihood of revealing reliable experimental effects. In comparison to ERPs evoked by visually presented words, those evoked by words presented auditorily may be associated with more latency jitter. This is due to the fact that, for auditorily presented words, there is more inter-item variation in the time after stimulus onset at which the stimulus can be recognised as a particular word. The principal factors which influence this variation are the speech rate, and the so-called uniqueness point of the individual words - the point in the word at which it can be uniquely identified (see Marslen-Wilson, 1984).

It is reasonable to assume that the additional variation in identification time would also result in increased variation in any subsequent processes indexed by the auditorily evoked ERPs. The principal effects of this additional variation would be to reduce the magnitude of any differences between experimental conditions, and, if more than one process is indexed by the ERPs, to increase the extent to which these processes overlap in the averaged ERP waveforms, thereby reducing the likelihood of observing any reliable changes in the distribution of experimental effects over time.

In the light of these observations a further study was conducted using the same general experimental paradigm employed in experiments 1 and 2. The modality manipulation was replaced by a voice manipulation, in order to ensure that recollection of the prior episode was the only basis upon which the contextual judgement could be made. At study subjects heard words, half of which were spoken in a male voice and half in a female voice. A test phase followed in which subjects made old/new and subsequent voice judgements to visually presented words.

7.2 *Methods*

Subjects: 18 subjects participated in the experiment, for which they were paid £3.50/hour. The data from two subjects were discarded due to excessive EOG artifact. Of the remaining 16 subjects, 7 were female, and all were right-handed as defined by writing hand. Subjects ages ranged from 18 to 23 years (average age 21).

Experimental Material: Stimuli consisted of 360 words and 90 pronounceable non-words. The words were divided into four lists, each list comprising 90 words. The sole difference between this experiment and experiment 1 with respect to list construction was the length of study and test lists. Study lists consisted of 270 critical items, divided into 3 equal blocks. Each block consisted of 60 words and 30 non-words. Half of these (30 words and 15 non-words) were spoken in the male voice, and half were spoken in the female voice. Each test list consisted of 360 words, half of which had been heard at study, and half of which were presented at test for the first time. Of the old words an equal number had previously been spoken in the male or the female voice. Each test list consisted of 6 blocks of 60 words, each block consisting of the same number of old and new words. The mappings between study and test lists were as for experiment 1, and the same number of study and test lists were produced (4 and 8 respectively). The characteristics of stimuli presented visually or auditorily were the same as for experiment 1, with the exception that the mean duration of auditorily presented stimuli was 620 msec.

Procedure: Subjects were exposed to one study list. Following a 5 minute delay they were exposed to one of the two test lists which corresponded to the list encountered at study. The study phase consisted of a modified lexical decision task. Following electrode placement (see general methods), subjects were seated in front of a TV monitor with the index finger and middle finger of each hand resting on microswitches. An asterisk preceded each trial and was removed 100 msec before the presentation of the stimulus. Subjects were instructed to respond to each item by pressing one of the four keys in front of them, depending upon whether the item was a word or a non-word, and whether the item had been spoken in the male or the female voice. This task differed from that in experiments 1 and 2, where only a lexical decision judgement was made. The intention of this modified lexical decision task was to orient subjects to speaker voice in order to encourage better memory for voice in the subsequent test phase.

For each subject the lexical decisions to items spoken in one of the two voices were always made with the two response keys on the same hand, whilst the lexical decisions to words spoken in the other voice were made with the alternate hand. The hands used for voice and lexical decision judgements were counter-balanced across subjects. Subjects were informed that accuracy and speed were of equal importance, and the fact that a recognition memory test would follow was not mentioned. A practice phase of 12 items preceded the study phase proper.

At test the procedure and stimulus presentation sequence was identical to that of experiment 1, with the exception of the fact that the modality judgement was replaced by a voice judgement (male vs female).

EEG Recording: EEG was recorded from the sites comprising the standard montage. EOG recording procedures and amplifier characteristics were as for experiment 1. On-line sampling was at 6 msec per point for a duration of 1536 msec, commencing 102 msec prior to stimulus onset.

7.3 *Results*

7.31 Behavioural data: Study phase

The probability of correct study word identification was 0.90 for words spoken in either the male or the female voice (male s.d. = .11, female s.d. = .09). The probabilities of correct non-word identification were 0.86 (s.d. = .10) for the male voice, and 0.84 (s.d. = .11) for the female voice. ANOVA on the behavioural data employed the factors of study voice (male vs female) and item type (word vs non-word). The analysis revealed that words attracted more correct judgements than did non-words ($F(1,15) = 12.45$; $p < .01$). Analysis of the study phase reaction times also employed the factors of study voice and item type. The analysis revealed that RTs to words were faster than RTs to non-words (1020 vs 1165 msec, $F(1,15) = 72.94$; $p < .001$).

7.32 *Behavioural data: Test phase*

Table 7.1 displays the probability of a correct old/new response for old and new test words, separated according to study voice. For words spoken in either voice, discrimination was above chance (for male $t(15) = 7.59$; $p < .001$, for female $t(15) = 12.31$; $p < .001$). Comparison of these discrimination measures revealed that they were not reliably different.

ANOVA on the RTs for old and new words (table 7.2) employed the factors of response accuracy (correct vs incorrect) and word type (male vs female vs new). The analysis revealed a main effect of both factors (respectively, $F(1.0,15.0) = 31.76$; $p < .001$, $F(1.9,29.1) = 3.46$; $p < .05$). The main effect of accuracy reflected the fact that correct responses were faster than incorrect responses. *Post-hoc* analyses (Newman-Keuls) revealed no reliable differences between the means for old female, old male, and new words. The largest RT differences are between the RTs to old female words and new words (1185 msec vs 1230 msec).

Table 7.3 displays the probability of a correct voice judgement for words spoken in either voice at study and correctly judged old at test. The probability of a correct voice judgement was 0.65, a value significantly above the chance probability of 0.50 ($t(15) = 5.56$; $p < .01$). The probability of a male voice judgement to a new word incorrectly judged old was 0.49. This value did not differ significantly from 0.50, suggesting that there was no response bias associated with voice judgements.

The RTs to correct and incorrect voice judgements are displayed in table 7.4. The RTs are separated according to the accuracy of the subsequent voice judgement. Analysis of these RTs employed the factors of response accuracy (correct vs incorrect) and word type (male vs female). No significant differences were revealed by the analysis.

7.33 *ERP Analyses*

A preliminary comparison of the hit/hit ERPs separated according to study voice revealed no significant differences. Appendix 2.1 displays the male hit/hit, female hit/hit and correct rejection ERPs. There were insufficient trials to permit a comparison of the hit/miss ERPs separated according to study voice. In all of the analyses that follow the hit/hit and hit/miss ERPs are collapsed across study voice. These collapsed ERPs are shown in figure 7.1, plotted against the ERPs to correct rejections. Following the initial N1 and P2 deflections, the waveforms consist of two principal late deflections. The first is negative and peaks approximately 500 msec post-stimulus. The second deflection is positive and peaks approximately 800 msec post-stimulus (later at anterior sites).

Figure 7.1 shows that at posterior electrode sites the hit/miss ERPs are more positive than the hit/hit and correct rejection ERPs. These differences are restricted to a small latency range centered around 300 msec post-stimulus. From approximately 400 msec post-stimulus both the hit/hit and hit/miss ERPs are more positive than the ERPs to correct rejections. For the hit/hit ERPs this positivity is maintained for the duration of the recording epoch at anterior sites, whilst the positivity associated with the hit/miss ERPs is more temporally restricted, diminishing within 300-500 msec of onset (700-900 msec post-stimulus). Post-1000 msec the magnitude of the differences between the hit/hit and correct rejection ERPs is greater at right rather than left frontal sites.

The ERPs shown in figure 7.1 were investigated by three planned analyses. The first two analyses compared the ERPs to correct rejections with the hit/hit and hit/miss ERPs respectively. The third comparison investigated the differences between the hit/hit and hit/miss ERPs. These analyses were performed over three latency regions: 500-800 msec, 800-1100 msec, and 1100-1400 msec. These regions overlap with those employed in previous ERP studies of recognition memory to visually presented test stimuli (Paller and Kutas, 1992; Rugg and Doyle, 1992), and with those employed in experiment 2, where a recording epoch of the same duration (1536 msec) was employed.

The results of these planned comparisons are displayed in table 7.5, which tabulates all effects involving response category. For the three latency regions analysed, appendix 1.3 displays mean amplitude measures for the three critical response categories at each electrode site.

7.331 *Analysis of hit/hit and correct rejection ERPs*

500-800 msec: Comparison of these ERPs at the midline revealed a main effect of response category and a response category x site interaction. As can be seen in figure 7.1, compared to the ERPs to correct rejections the midline hit/hit ERPs are more positive at Fz than at Cz and Pz respectively. A Scheffé analysis revealed that the differences between the hit/hit and correct rejection ERPs were reliably larger at Fz than at Pz.

The lateral analysis revealed a main effect of response category, and a response category x site interaction. These results in part reflect the fact that the differences between these ERPs are negligible at occipital sites, whilst at parietal and anterior sites the hit/hit ERPs are more positive. A Scheffé analysis comparing the average differences between the hit/hit and correct rejection ERPs at frontal and parietal sites with the differences between these ERPs at occipital sites revealed that the differences were reliably smaller at the occipital locations.

In addition to these anterior-posterior differences, at parietal sites the differences between these ERPs are larger at left than at right parietal locations. Given that previous reports of the parietal old/new effect have shown a left-greater than right parietal asymmetry (Neville *et al.*, 1986; Rugg *et al.*, 1995; Rugg and Doyle, 1992), a planned comparison of the differences between the hit/hit and correct rejection ERPs at LP and RP was performed, following the procedure employed in experiment 1. The analysis revealed a main effect of response category ($F(1,15) = 9.83$; $p < .01$), and a response category x site interaction ($F(1,15) = 8.24$; $p < .05$), reflecting the fact that the differences between these ERPs are larger at LP than at RP.

800-1100 msec: Comparison of the hit/hit and correct rejection ERPs revealed a main effect of response category at lateral sites, reflecting the fact that the hit/hit ERPs are more positive. The analyses also revealed response category x site interactions at midline and lateral sites. As can be seen in figure 7.1, these findings reflect the fact that the differences between these ERPs are largest at frontal electrode sites.

1100-1400 msec: Comparison of the hit/hit and correct rejection ERPs revealed response category x site interactions at midline and lateral sites, and a three-way interaction between response category, hemisphere and site. These results reflect the fact that compared to the ERPs to correct rejections the hit/hit ERPs are more positive at frontal sites, whilst at parietal and occipital sites these ERPs differ little, with the exception of Pz, where the hit/hit ERPs are more negative. The hemispheric asymmetry reflected in the response category x hemisphere x site interaction is most marked at frontal electrode locations. The magnitude of the differences between these ERPs is 3.2 μV at RF and 1.1 μV at LF (see appendix 1.3), although post-hoc analyses did not reveal a reliable difference in the size of the effects at these sites.

7.332 *Analysis of hit/miss and correct rejection ERPs*:

Table 7.5 shows that analysis of these ERPs revealed main effects of response category at midline and lateral sites over the 500-800 msec epoch, reflecting the fact that the hit/miss ERPs are more positive than the ERPs to correct rejections. Following the procedure employed for the analysis of the hit/hit old/new effects, a planned comparison of the hit/miss and correct rejection ERPs at the left and right parietal electrode sites was performed. The analysis revealed that the hit/miss ERPs are more positive ($F(1,15) = 20.11$; $p < .001$), but the interaction between category and site was not significant.

Over the 800-1100 msec epoch the hit/miss and correct rejection ERPs were not reliably different, whilst over the 1100-1400 msec epoch the only significant effect involving response category was an interaction between response category and site at the midline. This result reflects the fact that, as can be seen in figure 7.1, these ERPs differ little at Fz and Cz, but at Pz the hit/miss ERPs are more negative.

7.333 *Analysis of hit/hit and hit/miss ERPs*

Comparison of these ERPs revealed interactions between response category and site over the three latency regions analysed, at both midline and lateral sites. These interactions reflect the fact that in comparison to the ERPs to the hit/miss response category the hit/hit ERPs are relatively more positive at frontal sites, with the magnitude of the differences between these ERPs diminishing along the anterior-posterior axis, as can be seen in appendix 1.3. A planned comparison of the hit/hit and hit/miss ERPs at parietal sites revealed no effects involving response category.

7.334 *Analysis of onset latencies*

Three analyses of onset latencies were performed, investigating the time at which the ERPs to correct and incorrect voice judgements, and the ERPs to correctly identified new words, started to diverge.

The earliest reliable differences between the hit/hit and correct rejection ERPs occurred at 282 and 288 msec post-stimulus at sites Fz and RF respectively. The differences between the hit/miss and correct rejection ERPs onset 300 msec post-stimulus at the Cz electrode site. The earliest reliable differences between the hit/hit and hit/miss ERPs occurred at the posterior sites Pz, O1, and O2 at 258 msec post-stimulus. Following this early differentiation between these ERPs the next reliable onset of differences was at 540 msec at Fz.

These early posterior differences between the hit/hit, hit/miss and correct rejection ERPs were investigated by a one way analysis of variance comparing the ERPs over the 300-360 msec time window at sites O1 and O2. This time window straddles the latency which appears to maximally differentiate these ERPs. The analysis revealed a main effect of response category ($F(1.9,28.2) = 8.54$; $p < .001$). Whilst *post-hoc* analyses (Newman Keuls) revealed no reliable differences, the mean amplitudes for the three response categories were 2.54 μV (correct rejection), 2.10 μV (hit/hit), and 3.65 μV (hit/miss) respectively.

7.335 *Topographic analyses*

The hit/hit, hit/miss and correct rejection ERPs were further compared by two topographic analyses. First, the hit/hit and hit/miss ERPs were compared over the 500-800 msec time window - the epoch where both response categories were reliably different from the ERPs to correct rejections. The analysis revealed no reliable differences in scalp topography between these ERPs.

The second analysis investigated changes in the distribution of the differences between the hit/hit and correct rejection ERPs over the 500-800 and 1100-1400 msec epochs. Over the former epoch the differences between these ERPs are larger at left compared to right parietal sites, whilst at frontal sites the differences between these ERPs appear to be more symmetrical. Over the latter epoch there is little evidence for a parietal hemispheric asymmetry, but the ERPs differ more at right frontal compared to left frontal sites. These findings suggest a change in the distribution of the differences between these ERPs over time. This impression was confirmed by an analysis performed on the difference waveforms obtained by subtracting the correct rejection from the hit/hit ERPs, which revealed an interaction between epoch and site ($F(3.4,50.3) = 3.99$; $p < .01$).

The principal differences underlying this interaction were investigated by a subsidiary ANOVA on the right and left frontal and parietal electrode locations across the 500-800 and 1100-1400 msec epochs. The analysis revealed an interaction between epoch and hemisphere $F(1,15) = 6.04$; $p < .05$), reflecting the shift in distribution from a left-greater than right asymmetry over the earlier epoch to the reverse asymmetry over the later epoch. The rescaled values on which these analyses were performed can be seen in figure 7.2, which plots the relative values for the frontal and parietal electrode sites over the 500-800 and 1100-1400 msec epochs.

7.336 Analysis of misses and false alarms

The ERPs to misses, false alarms and correct rejections are shown in figure 7.3. The ERPs to misses and false alarms were each compared to the ERPs to correct rejections over the same epochs employed for the analysis of the hit/hit and hit/miss ERPs. Comparison of the ERPs to misses and to correct rejections revealed no effects involving response category across any epoch at either midline or lateral sites. Comparison of the ERPs to correct rejections and to false alarms revealed no evidence for positive-going effects similar to those observed for the analyses involving the hit/hit and hit/miss ERPs.

7.4 *Discussion*

ERPs evoked by words correctly assigned to study voice were more positive than those to correctly classified new words. These differences onset 300-400 msec post-stimulus, and continued, most markedly at frontal locations, until the end of the recording epoch. The distribution of the differences between these ERPs changed over time, suggesting that distinct processes contribute to memory for study context.

The differences between these ERPs from 500-800 msec are characterised by the left-greater-than-right parietal asymmetry reported in previous ERP studies of recognition memory (Neville *et al.*, 1986; Rugg and Doyle, 1992). Later in the recording epoch this parietal asymmetry is not evident, but the differences between the hit/hit and correct rejection ERPs are largest at frontal locations, and there is some evidence for a right-greater-than-left frontal asymmetry (for a report of a similarly distributed effect see Johnson, 1995). For ease of reference the earlier parietally distributed differences between the ERPs to correctly recognised old and new words will hereafter be referred to as the *parietal old/new effect*. The frontally distributed effect which is more extended in time will be referred to as the *frontal old/new effect*.

The ERPs evoked by words incorrectly assigned to study voice also displayed a parietal old/new effect, but no reliable frontal old/new effect. Over the 500-800 msec epoch the scalp distributions of the ERPs to correct and incorrect voice judgements were statistically indistinguishable, consistent with the view that the same processes were engaged in the two cases. Coupled with the findings that the hit/hit ERPs are characterised by both a parietal old/new effect and a frontal old/new effect, these findings are consistent with the view that recognition with and without retrieval of context share a common process, and contextual retrieval in addition depends upon the contribution of a second process (Moscovitch, 1992; Squire, 1994).

However, the data do not force the conclusion that correctly recognised words which were incorrectly assigned to context are not associated with neural activity in those regions which contribute to the frontal old/new effect. The absence of a frontal old/new effect for this response category may reflect the fact that the neural activity contributing to such an effect was too weak to be reliably detected. Consequently, the findings are consistent with the conclusions drawn on the basis of experiments 1 and 2: that the distinction between recognition with and without retrieval of context is one of degree rather than one of kind.

7.41 *The parietal old/new effect and relative fluency*

Given that the parietal old/new effect was reliable for words correctly and incorrectly assigned to study context, this effect is a candidate for an ERP index of processes related to fluency-based recognition. The data however provide little support for this view, since from 600 msec onwards the hit/hit ERPs tended to be more positive than the hit/miss ERPs at the left parietal site, although this difference was not statistically reliable. As for experiment 2, the data are therefore consistent with the view that ERPs are sensitive to relative fluency only if a relationship of redundancy holds between the processes of recollection and fluency-based recognition.

7.42 *Early differentiation of the hit/hit and hit/miss ERPs at occipital sites*

Whilst the view that the parietal old/new effect indexes fluency is only weakly supported by the data, the ERP analyses revealed an earlier latency region over which the differences between the hit/hit and hit/miss ERPs may be related to relative fluency: a restricted analysis of the hit/hit, hit/miss and correct rejection ERPs at the occipital sites O1 and O2 revealed a 60 msec time window centered on 330 msec post-stimulus where the hit/miss ERPs were more positive than either the hit/hit or the correct rejection ERPs.

These early differences between the ERPs to correct and incorrect voice judgements and those to correctly classified new words are therefore candidate indices of processes related to fluency-based recognition. Their early onset is consistent with the view that fluency-based recognition is related to facilitations in perceptual processing (Jacoby and

Dallas, 1981; Jacoby and Kelley, 1992). Further, the posterior distribution of these differences is at least suggestive of a posterior locus for the generators of the scalp-recorded activity. This is consistent with the findings in a recent PET study in which visual word form priming was linked to neural activity in right occipital cortex (Squire, Ojemann, Miezin, Petersen, Videen and Raichle, 1992). Two recent neuropsychological studies also describe impaired priming in a patient with right occipital lesions (Fleischman, Gabrieli, Reminger, Rinaldi and Morrell, 1995; Gabrieli, Fleischman, Keane, Reminger and Morrell, 1995).

If these early differences between the hit/hit, hit/miss and correct rejection ERPs do in fact reflect processes related to relative fluency, then the data suggest that two processes - indexed by these early effects and the parietal old/new effect respectively - are engaged on these types of task, both of which are related to recognition which is unaccompanied by retrieval of contextual information. Whilst previous discussions have focused on the differences between the respective frameworks of Jacoby and colleagues and Squire/Moscovitch (Jacoby and Kelley, 1992; Moscovitch, 1994; Squire and Knowlton, 1994), it is of course possible that both views are in part correct. That is, retrieval of information from declarative memory *and* relative fluency can support recognition memory judgements. This suggestion will be returned to in the general discussion (Chapter 11).

7.5 *Summary*

The differences between the ERPs to correct rejections and those to correctly recognised words which attracted correct context judgements were characterised by two principal modulations: a left-greater-than right parietal old/new effect, and a right-greater-than-left frontal old/new effect which was more extended in time. The ERPs to correctly recognised words which attracted an incorrect context judgement displayed a parietal old/new effect but no reliable frontal effect. These findings are consistent with the view that two processes contribute to memory for context, only one of which is necessary for recognition (Moscovitch, 1992; Moscovitch, 1994; Squire and Knowlton, 1994; Squire and Zola-Morgan, 1988). However, the findings are also consistent with the interpretation that the distinction between recognition with and without retrieval of context is quantitative rather than qualitative. This interpretation obtains because the absence of a frontal old/new effect for the hit/miss ERPs may reflect the fact that the ERP recordings were not sufficiently sensitive to reveal the neural activity differentiating the hit/miss and correct rejection ERPs at frontal electrode sites.

Table 7.1 Probability of correct old/new judgements to old and new words in experiment 3 (s.d. in brackets)

	Word Type		
	Female	Male	New
P(Correct Judgement)	0.71(0.10)	0.70(0.12)	0.72(0.04)

Table 7.2 Reaction times (msec) for correct and incorrect old/new judgements to old and new test words in experiment 3. Old words are separated according to speaker voice.

	Response	Word Type		
		Female	Male	New
RT	Correct	1103	1139	1201
SD		332	356	373
RT	Incorrect	1268	1263	1261
SD		426	399	386

Table 7.3 Conditional probability of a correct voice judgement to words correctly judged old in experiment 3. Also displayed (far right column) is the probability of a male voice judgement to a false alarm. (s.d. in brackets)

	Word Type		
	Female	Male	New
P(Correct Judgement)	0.64(0.14)	0.65(0.11)	0.49(0.11)

Table 7.4 Reaction times (msec) for initial old/new judgements in experiment 3, separated according to the accuracy of the subsequent voice judgement.

	Response	Word Type	
		Female	Male
RT	Correct	1079	1140
SD		316	357
RT	Incorrect	1145	1129
SD		336	341

Table 7.5 Results of pairwise analyses of the hit/hit, hit/miss, and correct rejection ERPs in experiment 3. The analyses were performed over the 500-800, 800-1100, and 1100-1400 msec epochs.

	500-800 msec				800-1100 msec				1100-1400 msec			
Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category (1,15)	12.40	10.27		**0.003**	0.12	13.11		n.s.	0.01	12.92		n.s.
Category x Site (2,30)	7.60	1.97	0.57	**0.012**	11.67	1.91	0.70	**0.001**	26.22	1.89	0.62	**0.001**
Lateral												
Category (1,15)	11.53	20.28		**0.004**	6.19	14.31		**0.025**	1.52	14.65		n.s.
Category x Site (4,60)	8.10	1.58	0.38	**0.004**	4.31	2.04	0.40	**0.033**	11.93	1.74	0.34	**0.001**
Category x Hem (1,15)	3.57	2.03		0.078	0.17	6.19		n.s.	2.45	2.56		n.s.
Category x Hem x Site (4,60)	3.13	0.51	0.54	0.053	0.72	1.68	0.40	n.s.	5.47	0.73	0.52	**0.009**
Hit/Miss vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	6.35	5.23		**0.024**	2.71	6.44		n.s.	0.06	13.04		n.s.
Category x Site	0.16	1.20	0.64	n.s.	0.39	1.22	0.66	n.s.	4.29	1.90	0.67	**0.041**
Lateral												
Category	21.69	3.81		**0.001**	0.17	7.99		n.s.	0.54	20.98		n.s.
Category x Site	0.99	1.18	0.48	n.s.	1.53	1.65	0.55	n.s.	0.81	2.23	0.38	n.s.
Category x Hem	1.20	3.55		n.s.	0.36	7.77		n.s.	0.09	3.16		n.s.
Category x Hem x Site	0.95	0.52	0.63	n.s.	0.36	1.82	0.42	n.s.	1.87	1.01	0.55	n.s.
Hit/Hit vs Hit/Miss	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	3.40	8.93		0.085	4.15	7.15		0.060	0.10	10.71		n.s.
Category x Site	11.11	1.60	0.54	**0.004**	9.38	1.77	0.58	**0.005**	11.72	1.52	0.61	**0.002**
Lateral												
Category	1.58	24.36		n.s.	2.95	23.13		n.s.	0.07	27.81		n.s.
Category x Site	9.71	1.45	0.40	**0.001**	5.62	1.87	0.50	**0.009**	7.68	1.54	0.46	**0.003**
Category x Hem	0.17	2.37		n.s.	0.22	3.06		n.s.	1.37	2.82		n.s.
Category x Hem x Site	1.04	0.40	0.50	n.s.	1.22	0.69	0.62	n.s.	1.19	0.59	0.55	n.s.

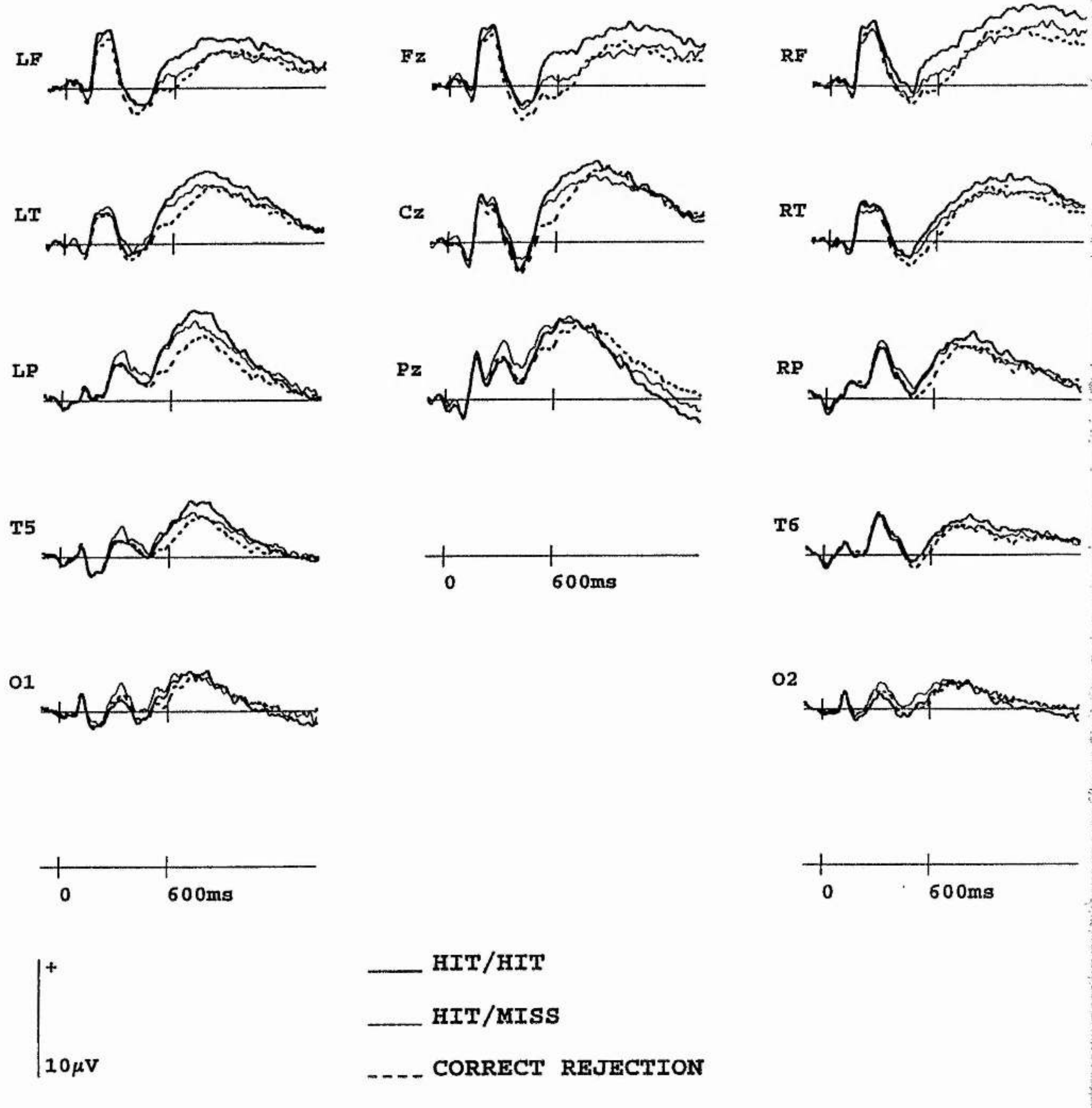
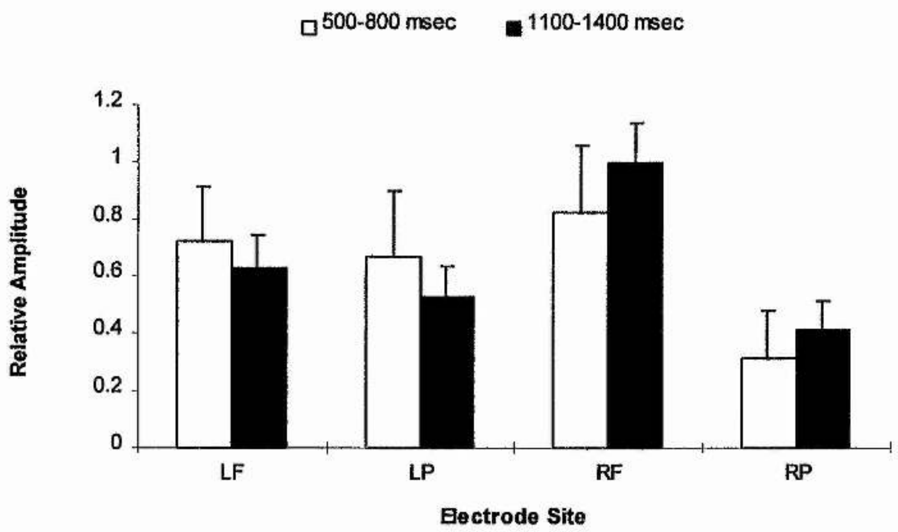


Figure 7.1 Grand average ERPs associated with the hit/hit, hit/miss, and correct rejection response categories in experiment 3. Electrode sites as for figure 5.1.

Figure 7.2 Rescaled mean amplitudes for the differences between the hit/hit and correct rejection ERPs in experiment 3. The relative amplitudes are shown for the left and right frontal (LF, RF) and parietal sites (LP, RP) over the 500-800 and 1100-1400 msec epochs.



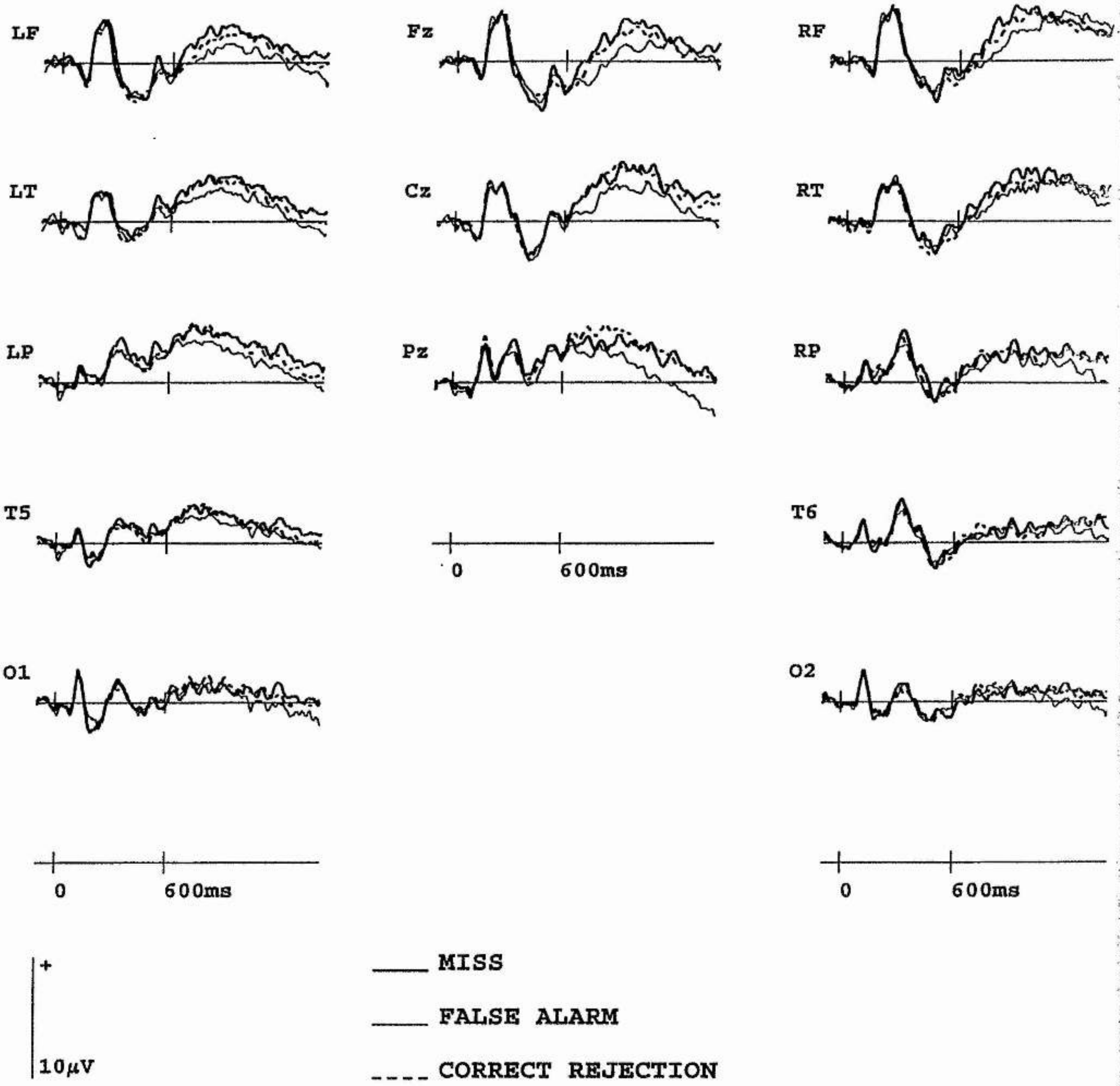


Figure 7.3 Grand average ERPs associated with the miss, false alarm, and correct rejection response categories in experiment 3. Electrode sites as for figure 5.1.

Chapter 8

8 *An ERP study of memory for words and memory for speaker voice: part two*

8.1 *Introduction*

In the design of the three experiments reported in the thesis to this point, both the initial old/new judgements and the subsequent context judgements have been forced choice. Consequently, there will have been a unknown proportion of trials on which subjects guessed when in possession of insufficient information to make an accurate judgement. The ERPs associated with such guesses will therefore have influenced the pattern of the ERP old/new effects reported in experiments 1-3.

For the initial old/new judgement a proportion of studied words will have been correctly guessed old. Since it is reasonable to assume that these 'correct' guesses will also be subject to a guess on the subsequent context judgement, they will therefore be distributed in some ratio between the ERPs to correct and incorrect context judgements. Given that a smaller number of trials comprise the ERPs to incorrect context judgements,[3] this response category will be influenced to the greater extent by such guesses.

[3] If this were not the case there would be no evidence that subjects could in fact make context discriminations at above chance level.

For the context judgement, if the initial old/new judgement was made on the basis of processes which did not entail retrieval of study context, then the subsequent context judgement is necessarily a guess. Consequently, the ERPs to correct context judgements will contain a proportion of trials which were not in fact accompanied by retrieval of contextual information. This proportion can be estimated, since in the absence of evidence for any systematic bias for context judgements, the proportion of incorrect voice judgements (for words correctly judged old) represents half the total proportion of context judgements made without retrieval of veridical contextual information. In experiment 3 the conditional probability of an incorrect context judgement was 0.35. Hence approximately 50% (0.35/(1 - 0.35)) of the trials comprising the ERPs to correct context judgements may not in fact have been associated with retrieval of accurate contextual information.

The likely consequence of these guesses, on both the first and second test judgements, would be to reduce the extent to which any hit/hit and hit/miss ERP old/new effects can be considered accurate representations of ERP signatures of recognition with and without retrieval of study context. This is critical, since the principal theoretical focus of these experiments is on the differences between the old/new effects associated with the hit/hit and the hit/miss response categories.

Experiment 4 attempted to circumvent some of the potential problems associated with the forced choice nature of experiments 1-3 by introducing a 'don't know' option for the old/new judgement and the subsequent context judgement. The intention of this manipulation was to investigate the extent to which the old/new effects reported in experiment 3 were influenced by the distribution of 'correct' guesses in the ERPs to the hit/hit, hit/miss and correct rejection response categories.

8.2 *Methods*

Subjects: 19 subjects took part in the experiment, for which they were paid £3.50/hour. The data from 2 subjects were discarded due to a technical error. The data from a further subject was discarded because too few correct old judgements were made to permit formation of reliable averaged waveforms when these trials were separated according to the accuracy of the subsequent voice judgement. Of the remaining 16 subjects all were right handed, and 12 were female. The average age of subjects was 19 years (range 17-24).

Experimental Material: The items were drawn from the same word and non-word pools as in previous experiments (see appendices 3.1 and 3.2). The same number of stimuli were employed as in experiment 3. Study and test lists were formed using the same procedure as outlined for experiments 1 to 3. Visual stimuli subtended a maximum horizontal visual angle of 2.0 degrees, and a maximum vertical angle of 0.6 degrees. Auditory stimuli were presented binaurally at a comfortable hearing level. They were digitised at 22 Khz with 16 bit resolution, and stored on the hard disk of an IBM-compatible PC. Mean duration for auditorily presented stimuli was 660 msec for words spoken in the male voice, and 630 msec for words spoken in the female voice.

Procedure: The only difference from the study phase procedure adopted in experiment 3 was that the inter-trial interval was lengthened to 4.1 seconds. The additional delay preceded presentation of the fixation asterisk on each trial. For the test phase, all aspects of the procedure were as for experiment 3, with the exception that subjects had the option to make a don't know response for both test judgements. As in experiment 1, the old/new and voice judgements were made on the response keys on which the index fingers of the subjects rested. Subjects made a don't know response by pressing one of the response keys on which their middle fingers rested: the finger used for this response was counterbalanced across subjects. The hands required for the first and second judgement were also counterbalanced across subjects so that there was no correlation between the old/new and male/female judgements.

EEG Recording: EEG and EOG recording procedures were as for experiment 3, with the following exceptions. First, EEG was recorded from an additional 4 electrode sites. These additional sites were left and right prefrontal (FP1, FP2), and the superior parietal sites (P3, P4). Second, all channels were amplified with a bandpass of 35 to 0.03 Hz (3 dB points). Third, during recording all channels were referenced to the left mastoid, and EEG was recorded from the right mastoid. The EEG from all channels was re-referenced offline to linked mastoids.

8.3 *Results*

8.31 *Behavioural data: Study phase*

For words spoken in the male or female voice, the probabilities of correct study word identification were 0.92 and 0.93 respectively (male s.d. = .11, female s.d. = .09). For non-words the respective probabilities were 0.86 (s.d. = .10), and 0.84 (s.d. = .11). ANOVA on the behavioural data employed the factors of study voice (male vs female) and item type (word vs non-word). The analysis revealed that words attracted more correct judgements than did non-words ($F(1.0,15.0) = 12.45$; $p < .01$). Analysis of the study phase reaction times employed the same factors. The analysis revealed that RTs to words were faster than RTs to non-words (1020 vs 1165 msec, $F(1.0,15.0) = 72.94$; $p < .001$), and that responses to items spoken in the female voice were faster than responses to items spoken in the male voice (1124 vs 1165 msec, $F(1,15) = 19.91$; $p < .001$). As noted in the methods section of this chapter, the average length of the female speech samples was 30 msec less than that of the male speech samples, which in all likelihood contributed to this difference in study phase RTs.

8.32 *Behavioural data: Test phase*

Table 8.1 displays the probability of correct, incorrect, and don't know judgements for the initial recognition decision to old and new test words. Old words are separated according to study voice. A discrimination measure of *p(hit) -(p(false alarm) + p(don't know))* was computed for words spoken in either voice. This measure represents the lower bound on discrimination estimates commonly obtained on tests of recognition memory by the index *p(hit) - p(false alarm)*.

For words spoken in either voice, discrimination assessed by this modified formula was above chance (male $t(15) = 7.01$; $p < .001$, female $t(15) = 8.10$; $p < .001$). Comparison of the two discrimination measures revealed no significant differences. ANOVA comparing the probabilities of incorrect responses to old female, old male, and new words also revealed no significant differences. However, ANOVA comparing the probabilities of a don't know response to these three word types revealed a main effect ($F(1.3,19.5) = 6.56$; $p < .05$). *Post-hoc* analyses (Newman Keuls) revealed that whilst the probabilities of don't know responses to old female and old male words were not reliably different, both were significantly lower than the probability of a don't know response to a new word.

Table 8.2 displays the RTs for correct, incorrect, and don't know judgements to old and new words. Given the low number of don't know responses, and the fact that two subjects made no don't know responses for the first decision, analysis of the RTs was restricted to correct and incorrect judgements. The analysis employed the factors of accuracy (correct vs incorrect) and word type (new vs old male vs old female). The analysis revealed a main effect of accuracy ($F(1.0,15.0) = 32.29$; $p < .001$), and an interaction between accuracy and word type ($F(1.4,20.3) = 8.35$; $p < .01$). *Post-hoc* analyses (Newman Keuls) revealed that whilst incorrect judgements were slower than correct judgements for old words, RTs to new words did not differ as a function of response accuracy. Further, correct responses to old words were faster than correct responses to new words, but the RTs for incorrect responses did not differ according to word type.

Table 8.3 displays the probabilities of male, female, and don't know judgements for words judged old. For words spoken in either the male or the female voice, the probability of a correct voice judgement was reliably higher than the probability of an incorrect judgement (respectively, $t(15) = 4.17$; $p < .001$, and $t(15) = 5.77$; $p < .001$). Comparison of the probabilities of male and female voice judgements to new words incorrectly judged old revealed no evidence for a voice response bias. ANOVA comparing the probabilities of don't know responses to old words (male and female) and to new words incorrectly judged old revealed a main effect of condition ($F(1.4,21.7) = 36.53$; $p < .001$). *Post-hoc* analyses (Newman Keuls) revealed that whereas the probabilities of a don't know judgement to old male and old female words did not differ, both were significantly lower than the probability of making a don't know response to a new word that had been incorrectly judged old.

The RTs for correct, incorrect, and don't know responses to male and female words are displayed in table 8.4. ANOVA on these RTs employed the factors of response accuracy (correct vs incorrect vs don't know) and word type (male vs female). The analysis revealed a main effect of response accuracy ($F(1.8,27.2) = 10.82$; $p < .001$). *Post-hoc* tests (Newman Keuls) revealed that whereas the RTs for correct and incorrect judgements did not differ, both were reliably faster than the RTs for don't know responses. Insufficient incorrect old judgements to new words were made to permit a reliable RT comparison when these words were separated according to the subsequent voice judgement.

8.33 *ERP Analyses*

As previously noted (general methods, chapter 4), the initial ERP analyses reported here are over the sites comprising the standard montage. ERP analyses including the additional sites from which ERPs were recorded in this experiment (FP1, FP2, P3, P4) revealed qualitatively similar results.

Trials associated with don't know responses for the initial old/new judgement were discarded. As for experiment 3 a preliminary analysis comparing the ERPs to correct voice judgements separated according to study voice revealed no reliable differences, and in all of the analyses below the ERPs to correct and incorrect voice judgements are collapsed across study voice. Of the 16 subjects included in the experiment only 12 made sufficient incorrect voice judgements to permit a comparison of the hit/miss and correct rejection ERPs. These hit/miss ERPs were not reliably different from those formed by pooling trials to correctly recognised items which were associated with incorrect and don't know voice judgements. The ERPs to these two response categories are displayed in appendix 2.2. Since collapsing across these response categories permitted analyses involving all 16 subjects, the analyses reported below will be for this collapsed category. This will be referred to as the *hit/miss* response category, whilst noting that the category is not strictly equivalent to the hit/miss category as defined in experiment 3.

Figure 8.1 displays the collapsed hit/hit, hit/miss and correct rejection ERPs, where it can be seen that both the hit/hit and hit/miss ERPs are more positive than the ERPs to correct rejections from approximately 400 msec post-stimulus. For the hit/miss ERPs this positivity diminishes after 300-400 msec. However, at the right frontal site the hit/miss ERPs are also more positive than the ERPs to correct rejections from 1100 msec onwards. From approximately 400 msec post-stimulus the hit/hit ERPs are also more positive than the hit/miss ERPs. Appendix 2.3 shows the ERPs to these three response categories for all 17 scalp sites from which EEG was recorded in this experiment.

The same analysis strategy was employed for these ERPs as was employed in experiment 3, for both the principal analyses of variance and the analyses of onset latencies. Table 8.5 displays the results of the three planned comparisons of the hit/hit, hit/miss and correct rejection ERPs over the 500-800, 800-1100, and 1100-1400 msec time windows. For these latency regions, appendix 1.4 displays mean amplitude measures for the three critical response categories for each electrode site comprising the standard montage.

8. 331 *Analysis of hit/hit and correct rejection ERPs*

500-800 msec: Comparison of these ERPs revealed main effects of response category at midline and lateral sites, reflecting the fact that the hit/hit ERPs are more positive. The analyses also revealed a response category x site interaction at the midline. Scheffé analyses revealed that the differences between the hit/hit and correct rejection ERPs were not reliably different at Fz and Cz, but the differences between the ERPs at these sites were reliably larger than those at Pz.

Following the procedure employed in experiment 3, a planned comparison of the hit/hit and correct rejection ERPs at the left and right parietal electrode sites was performed. The analysis revealed a main effect of response category ($F(1,15) = 12.73$; $p < .01$), and the interaction between category and site approached significance ($F(1,15) = 4.05$; $p = .06$). The sizes of the old/new effects for the hit/hit ERPs at LP and RP are 3.4 μV and 2.3 μV respectively.

800-1100 msec: Comparison of the hit/hit and correct rejection ERPs at lateral sites revealed a main effect of response category, and a response category x hemisphere x site interaction. These findings reflect the fact that the differences between these ERPs are relatively smaller at posterior locations than at frontal and parietal locations. In addition, at frontal sites the differences between these ERPs are larger over the right hemisphere than over the left (3.3 μV vs 2.0 μV), whereas at parietal sites the opposite asymmetry obtains (2.3μV vs 3.4μV). The analysis at the midline revealed an interaction between category and site, reflecting the fact that the differences between these ERPs are largest at frontal locations.

1100-1400 msec: Comparison of the hit/hit and correct rejection ERPs revealed response category x site interactions at midline and lateral sites, a response category x hemisphere interaction, and a three-way interaction between response category, hemisphere and site. These results in part reflect the fact that compared to the ERPs to correct rejections the hit/hit ERPs are more positive at frontal sites, whilst at parietal and occipital sites these ERPs differ little. The response category x hemisphere interaction reflects the fact that the differences between these ERPs are larger at right than at left hemisphere sites.

On the basis of the trend in experiment 3 for the differences between the hit/hit and correct rejection ERPs to be larger at the right frontal electrode site than at it's contralateral homologue, a planned comparison of these response categories at frontal locations was performed. The analysis revealed a main effect of response category ($F(1,15) = 6.30$; $p < .05$), and a response category x site interaction ($F(1,15) = 21.07$; $p < .001$), reflecting the fact that the differences between these ERPs are larger at the right than at the left frontal electrode site.

8.332 *Analysis of hit/miss and correct rejection ERPs*

500-800 msec: Table 8.5 shows that analysis of these ERPs revealed response category x site interactions at midline and lateral sites, and an interaction between these two factors and hemisphere. Figure 8.1 shows that at midline sites the differences between the hit/miss and correct rejection ERPs are largest at Fz, whilst at lateral sites the differences between these ERPs are smallest at occipital sites. A planned comparison of these ERPs restricted to the parietal sites revealed a significant interaction between response category and site ($F(1,15) = 6.23$; $p < .05$), reflecting the fact that the differences between these ERPs are larger at the left than at the right parietal site (1.9 μV vs 1.0 μV).

800-1100 msec: Comparison of the hit/miss and correct rejection ERPs revealed a response category x site interaction at the midline, and a response category x hemisphere x site interaction. These results reflect the fact that in comparison to the ERPs to correct rejections the hit/miss ERPs are relatively more positive at frontal locations than at posterior locations. In addition, at frontal locations the differences between these ERPs are more marked over the right hemisphere, whereas at parietal and posterior temporal sites the opposite asymmetry obtains (see appendix 1.4).

1100-1400 msec: Comparison of the hit/miss and correct rejection ERPs over this epoch revealed a response category x site interaction at midline and lateral sites, and a response category x hemisphere x site interaction. These results in part reflect the fact that the hit/miss ERPs are relatively more positive than the ERPs to correct rejections at anterior sites, and more negative at posterior sites. The most marked hemisphere asymmetry in the size of the differences between these ERPs is at the frontal sites, where the ERPs differ little at the left frontal site, but at the right frontal site the hit/miss ERPs are more positive. An analysis of the hit/miss old/new effects restricted to the frontal sites revealed an interaction between response category and site ($F(1,15) = 12.13$; $p < .01$), reflecting the fact that the hit/miss ERPs are relatively more positive than the ERPs to correct rejections at RF than at LF.

8.333 *Analysis of hit/hit and hit/miss ERPs*

Comparison of these ERPs revealed main effects of response category at midline and lateral sites across all three epochs, with the exception of the analysis at midline sites from 1100-1400 msec post-stimulus. These results reflect the fact that the hit/hit ERPs are more positive than the hit/miss ERPs. Over the 500-800 msec epoch a planned comparison of the hit/hit and hit/miss ERPs at the left and right parietal sites revealed a main effect of response category, reflecting the fact that the hit/hit ERPs are the more positive ($F(1,15) = 7.38$; $p < .05$).

8.334 *Analysis of onset latencies*

The onset latency analyses reported here are for those sites at which the earliest differences occurred, and for those sites reported in the analysis of onset latencies in the previous chapter. The earliest reliable differences between the hit/hit and correct rejection ERPs in this experiment occurred at 390 msec post-stimulus at Cz. For the Fz electrode site the earliest reliable differences onset 396 msec post-stimulus. The differences between the hit/miss and correct rejection ERPs onset 300 msec post-stimulus at the left-temporal electrode site, and 312 msec post-stimulus at the central midline electrode Cz. The differences between the hit/hit and hit/miss ERPs onset 558 msec post-stimulus at LP. The analyses comparing the differences between these ERPs at occipital sites did not reveal 10 significant and consecutive t-values over any latency region.

8.335 *Topographic analyses*

Differences in the scalp topography of the old/new effects for the hit/hit and hit/miss ERPs were investigated by a single ANOVA which compared the distributions of the hit/hit and hit/miss ERP old/new effects across the 500-800 and 1100-1400 msec epochs. The analysis revealed an interaction between epoch and electrode site ($F(5.3,78.9) = 5.08$; $p < .001$), indicating changes in scalp distribution over time.

This interaction was further investigated by a subsidiary ANOVA which was restricted to the frontal and parietal electrode sites. The analysis revealed an interaction between epoch and hemisphere ($F(1,15) = 29.67$; $p < .001$). The rescaled data values for these electrode sites over the two epochs are shown in figure 8.2, where it can be seen that the distribution shifts from a left-greater-than-right asymmetry over the 500-800 msec epoch to the opposite asymmetry over the 1100-1400 msec epoch. These results are qualitatively similar to those reported in experiment 3 (see figure 7.2), although figure 8.2 does suggest a more marked anterior shift in scalp distribution over time.

8.336 *Analysis of misses and false alarms*

The ERPs to misses, false alarms and correct rejections are shown in figure 8.3. The ERPs to misses and false alarms were each compared to the ERPs to correct rejections over the same epochs employed for the analysis of the hit/hit and hit/miss ERPs. Comparison of the ERPs to misses and to correct rejections revealed no effects involving response category across any epoch at either midline or lateral sites. Comparison of the ERPs to correct rejections and to false alarms revealed no evidence for positive-going effects similar to those observed for the analyses involving the hit/hit and hit/miss ERPs.

8.4 *Discussion*

The findings in this experiment differed from those of experiment 3 in three principal ways. First, there was no evidence for any early posterior differentiation between the ERPs to correct and incorrect voice judgements, and the ERPs to words correctly judged new. The absence of such differences provides no support for the view that between 250 and 350 msec post-stimulus the occipital electrode sites are sensitive to processes related to fluency-based recognition. However, the absence of these early differences in this experiment may stem in part from the introduction of the don't know response option. This option was introduced in order to attract low confidence judgements where the subject was uncertain of the old/new status of the test item. If this response option in fact attracted low confidence fluency-driven responses which, in the absence of the don't know option, would have been correctly judged old, then this would go some way to explaining the absence of the early differentiation between the hit/hit and hit/miss ERPs at posterior sites in this experiment.

Second, over the 1100-1400 msec epoch a statistically reliable frontal old/new effect was evident for the hit/miss response category. The distribution of this effect was indistinguishable from that of the frontal old/new effect for the hit/hit response category. There are two possible reasons why this frontal old/new effect was not found in experiment 3. The first explanation is that, as predicted, the don't know option for the old/new judgement effected a clearer distinction between the ERPs to correctly identified old and new words by reducing the proportion of 'correct' guesses contributing to these ERPs. The second explanation is that the introduction of the option to respond 'don't know' in this experiment attracted responses associated with weak recollection, which would have been correctly assigned to study context had the 'don't know' option been unavailable. However, in the 12 subjects who made sufficient incorrect voice judgements to permit the formation of reliable averaged waveforms, the comparison of the 'genuine' hit/miss ERPs and the collapsed hit/miss and 'don't know' ERPs revealed no reliable differences, suggesting that the reliable frontal old/new effect in experiment 4 was not due to the fact that the 'don't know' option attracted weakly recollected words (see figure in appendix 2.2).

The third difference between these findings and those of experiment 3 was that the differences between the hit/hit and hit/miss ERPs were larger in the present experiment. In particular, at parietal sites the hit/hit ERP old/new effect was reliably larger than the hit/miss old/new effect. This finding is inconsistent with the view that the parietal old/new effect is sensitive to fluency, irrespective of which model is assumed to hold between the processes of fluency and recollection (see chapter 3).

The larger old/new effects revealed in this experiment are most likely attributable to the introduction of the don't know options. Whereas the don't know option for the initial old/new judgement was introduced principally to reduce the impact of 'correct' guesses on the hit/miss old/new effect, the introduction of the option for the subsequent voice judgement was intended to reduce the proportion of trials contributing to the hit/hit ERPs which were not associated with veridical memory for context. There is some behavioural evidence that subjects employed the 'don't know' option for items which were associated with little or no accurate contextual information. This stems from the finding that subjects were twice as likely to make a 'don't know' judgement to new words incorrectly judged old as they were to make a male or female voice judgement to those same words.

Whilst the preceding paragraphs have discussed the differences between the findings across the two experiments, the critical aspects of the two experiments are their similarities, and the converging conclusions which the results permit. The principal similarities between the ERP effects revealed in experiments 3 and 4 are shown in figure 8.4, which plots, for frontal and parietal sites, the respective hit/hit, hit/miss and correct rejection ERPs.

8.41 *Comparison of onset latencies*

In both experiments the analyses of onset latencies revealed that the differentiation between the hit/hit and hit/miss ERPs was some 200 msec later than the time at which these two classes of ERPs diverged from the ERPs to correct rejections. These findings are consistent with reports that accurate recognition judgements can be made at shorter latencies than can accurate context judgements (Dosher, 1984; Hintzman and Curran, 1994; Johnson *et al.*, 1994). However, the findings do not necessarily indicate that contextual retrieval begins some time after retrieval of item information. Retrieval of these forms of information may occur in parallel, but the time at which sufficient information is available for an item judgement may precede that at which sufficient information is available to make a context judgement. Note that in experiment 3 there was an early differentiation between the hit/hit and hit/miss ERPs at occipital sites, however, as discussed above the reliability of these effects has not been established.

8.42 *The processing indexed by parietal and frontal old/new effects*

In this experiment both the hit/hit and hit/miss ERPs were characterised by topographically distinct parietal and frontal old/new effects. The only reliable statistical effects differentiating these ERPs were main effects of response category, emphasising the fact that the differences between these ERPs were purely quantitative. As for experiment 3, these findings are consistent with the view that while more than one process contributes to memory for context, these same processes are also engaged when recognition is not accompanied by the ability correctly to assign a test word to study context. The results therefore provide little support for the view that distinct processes contribute to recognition with and without retrieval of context (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992).

8.43 *The parietal old/new effect and retrieval from declarative memory*

It has been proposed that the parietal old/new effect is sensitive to processes related to recollection (Paller and Kutas, 1992; Paller *et al.*, 1995; Smith, 1993; Smith and Halgren, 1989), and that the size of the effect varies with the quality or amount of information retrieved from memory (Rugg *et al.*, 1995). Rugg and colleagues (1995) based this proposal on the finding that low-frequency words were associated with larger old/new effects than high- frequency words, even when both frequency classes were correctly assigned to the context in which they had been presented at study, and all judgements were rated by the subject as highly confident. They suggested that the differences between the old/new effects for low- and high- frequency words arose because low-frequency words engendered retrieval of more information at study than did high- frequency words.

The fact that in this experiment the hit/hit ERPs were associated with a reliably larger parietal old/new effect than were the hit/miss ERPs is consistent with the view that the parietal old/new effect is sensitive to the amount or the quality of information retrieved from memory. In experiment 3 the hit/hit ERPs tended to be larger at parietal sites, although the difference between these ERPs was not reliable. The preceding comments regarding the introduction of the don't know response option suggest that the absence of a reliable difference between these two classes of ERPs in experiment 3 may have been due to the attenuating influence of trials contributing to the hit/hit ERPs which were not associated with retrieval of accurate contextual information.

8.44 *The frontal old/new effect and retrieval from declarative memory*

In contrast to the parietal old/new effect, the frontal effect is maximal at frontal scalp locations and has a more extended time course. In experiments 3 and 4 this effect was reliably larger for the hit/hit than for the hit/miss ERPs, suggesting that it plays a functional role in context judgements. In this experiment there was also evidence for a reliable, albeit diminished, frontal effect in the hit/miss ERPs. These findings suggest that, like the parietal old/new effect, the frontal effect is a graded process, sensitive to the quality or amount of information retrieved from memory.

These functional considerations regarding the parietal and frontal old/new effects are therefore wholly consistent with the view that recognition with and without retrieval of context share a common retrieval function - indexed by the parietal old/new effect, and that an additional process - indexed by the frontal old/new effect -is required for the retrieved information to be placed in its correct context (Moscovitch, 1994; Squire and Zola-Morgan, 1988).

Table 8.1 Probabilities of correct, incorrect and don't know judgements to old and new words for the initial recognition judgement in experiment 4. (s.d. in brackets)

	Word Type		
	Female	Male	New
P(Correct)	0.70 (0.09)	0.67 (0.11)	0.62 (0.16)
P(Incorrect)	0.21 (0.05)	0.22 (0.09)	0.23 (0.13)
P(DON'T KNOW)	0.09 (0.13)	0.11 (0.15)	0.15 (0.15)

Table 8.2 Reaction times (msec) to correct, incorrect, and don't know judgements for the initial recognition judgement in experiment 4.

	Response	Word Type		
		Female	Male	New
RT	Correct	1206	1203	1382
SD		322	313	358
RT	Incorrect	1477	1514	1425
SD		341	337	369
RT	DON'T KNOW	1867	1810	1907
SD		229	305	276

Table 8.3 Conditional probabilities of correct, incorrect and don't know judgements to words correctly judged old in experiment 4. (s.d. in brackets)

	Word Type		
	Female	Male	New
P(Male)	0.27 (0.15)	0.50 (0.13)	0.23 (0.15)
P(Female)	0.50 (0.13)	0.24 (0.14)	0.26 (0.18)
P(DON'T KNOW)	0.23 (0.12)	0.26 (0.11)	0.51(0.18)

Table 8.4 Reaction times (msec) to words correctly judged old, separated according to the subsequent voice judgement (correct, incorrect, don't know).

	Response	Word Type	
		Female	Male
RT	Correct	1163	1160
SD		315	291
RT	Incorrect	1195	1233
SD		261	271
RT	DON'T KNOW	1291	1348
SD		298	314

Table 8.5 Results of the pairwise analyses of the hit/hit, hit/miss, and correct rejection ERPs in experiment 4. The analyses were performed over the 500-800, 800-1100, and 1100-1400 msec epochs.

	500-800 msec				800-1100 msec				1100-1400 msec			
Hit/Hit vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category (1,15)	10.20	20.67		**0.006**	4.07	25.62		0.062	1.08	27.03		n.s.
Category x Site (2,30)	6.58	0.55	0.64	**0.014**	5.05	2..46	0.62	**0.031**	16.03	2.15	0.57	**0.001**
Lateral												
Category (1,15)	12.52	36.87		**0.003**	9.69	39.06		**0.007**	1.56	57.47		n.s.
Category x Site (4,60)	3.05	1.55	0.44	0.070	2.56	3.96	0.34	n.s.	9.49	3.65	0.39	**0.002**
Category x Hem (1,15)	2.31	2.52		n.s.	0.09	4.95		n.s.	5.69	4.59		**0.031**
Category x Hem x Site (4,60)	2.79	0.53	0.53	0.074.	5.77	0.93	0.53	**0.007**	9.41	0.88	0.62	**0.001**
Hit/Miss vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	2.89	13.30		n.s.	0.08	12.17		n.s.	0.05	28.70		n.s.
Category x Site	9.32	0.54	0.70	**0.003**	7.80	1.09	0.72	**0.005**	18.29	1.17	0.81	**0.001**
Lateral												
Category	2.86	24.59		n.s.	0.15	16.58		n.s.	0.05	33.92		n.s.
Category x Site	3.55	0.89	0.52	**0.040**	1.52	1.96	0.39	n.s.	8.56	1.73	0.43	**0.002**
Category x Hem	2.86	1.79		n.s.	0.53	4.58		n.s.	1.17	6.36		n.s.
Category x Hem x Site	3.88	0.30	0.58	**0.025**	4.77	0.81	0.52	**0.015**	7.02	1.01	0.58	**0.002**
Hit/Hit vs Hit/Miss	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	5.20	13.31		**0.038**	14.51	8.61		**0.002**	2.72	16.05		n.s.
Category x Site	0.24	0.71	0.64	n.s.	0.67	1.48	0.62	n.s.	1.94	1.47	0.61	n.s.
Lateral												
Category	6.13	27.97		**0.026**	20.39	15.72		**0.001**	4.68	24.67		**0.047**
Category x Site	0.80	1.44	0.44	n.s.	1.28	2.83	0.39	n.s.	1.91	2.92	0.43	n.s.
Category x Hem	0.01	1.84		n.s.	0.31	2.55		n.s.	1.62	3.50		n.s.
Category x Hem x Site	1.02	0.47	0.50	n.s.	0.91	0.60	0.57	n.s.	0.95	0.75	0.51	n.s.

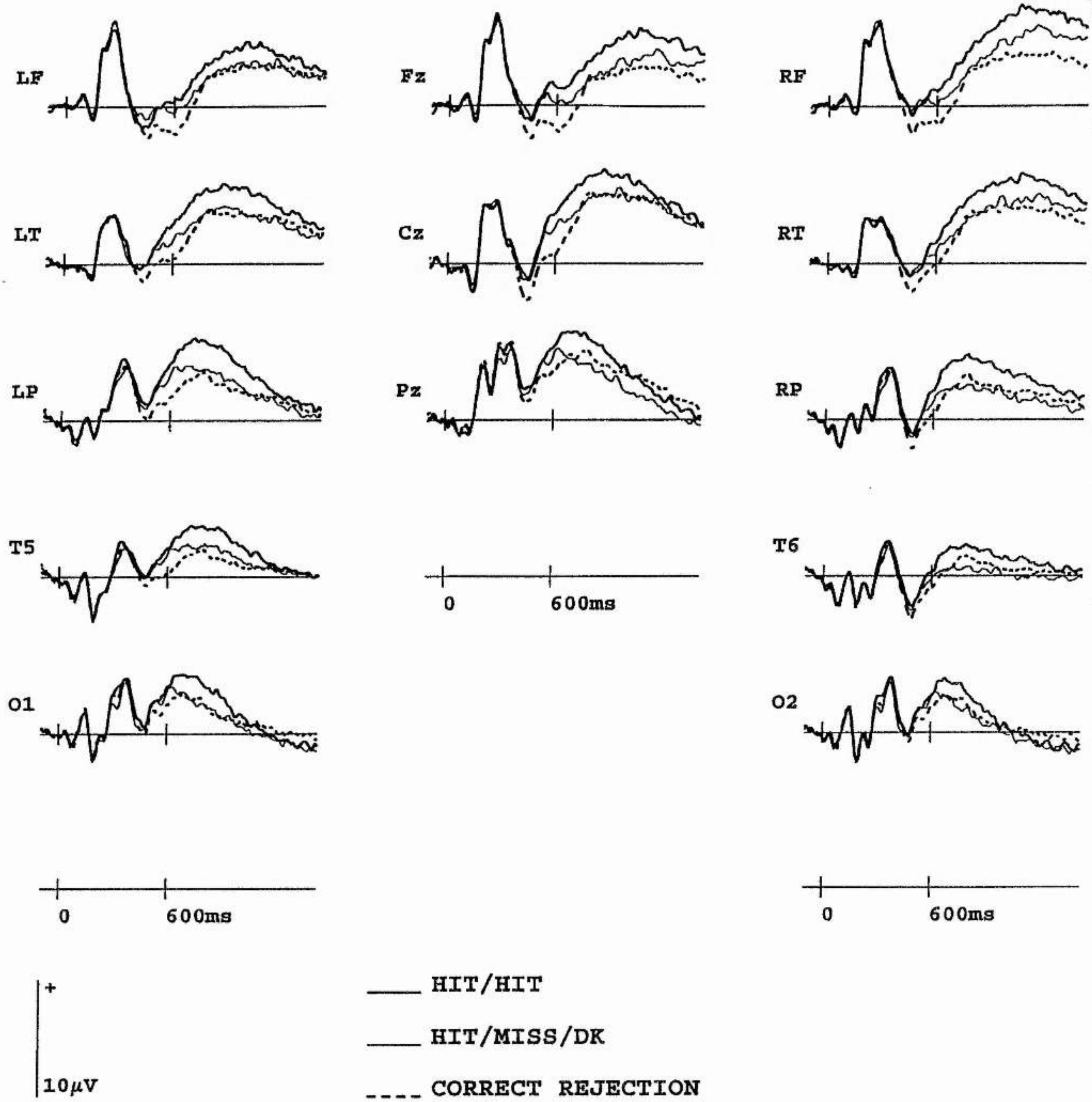
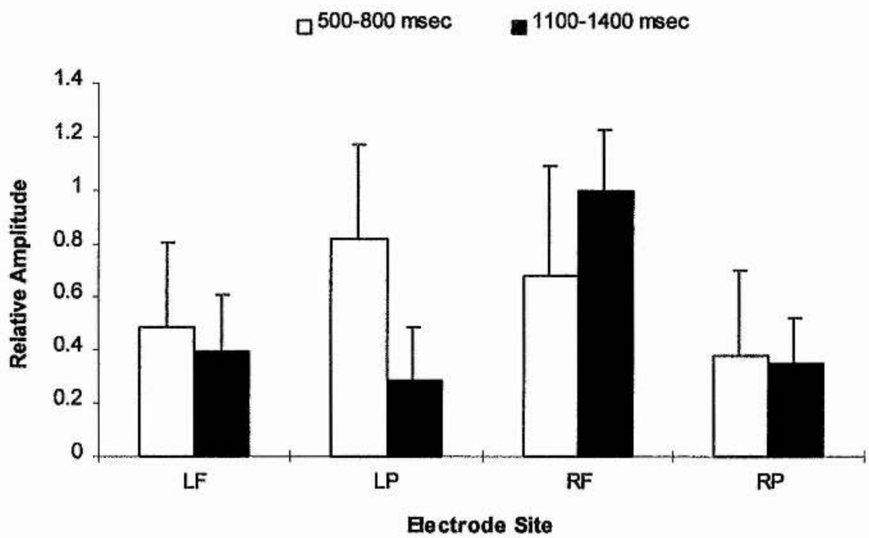


Figure 8.1 Grand average ERPs associated with the hit/hit, hit/miss/dk, and correct rejection response categories in experiment 4. Electrode sites as for figure 5.1.

Figure 8.2 Rescaled amplitudes of the differences between the ERPs to words correctly judged old and the ERPs to correct rejections in experiment 4. The relative amplitudes are shown for the left and right frontal (LF, RF) and parietal (LP, RP) sites over the 500-800 and 1100-1400 msec epochs.



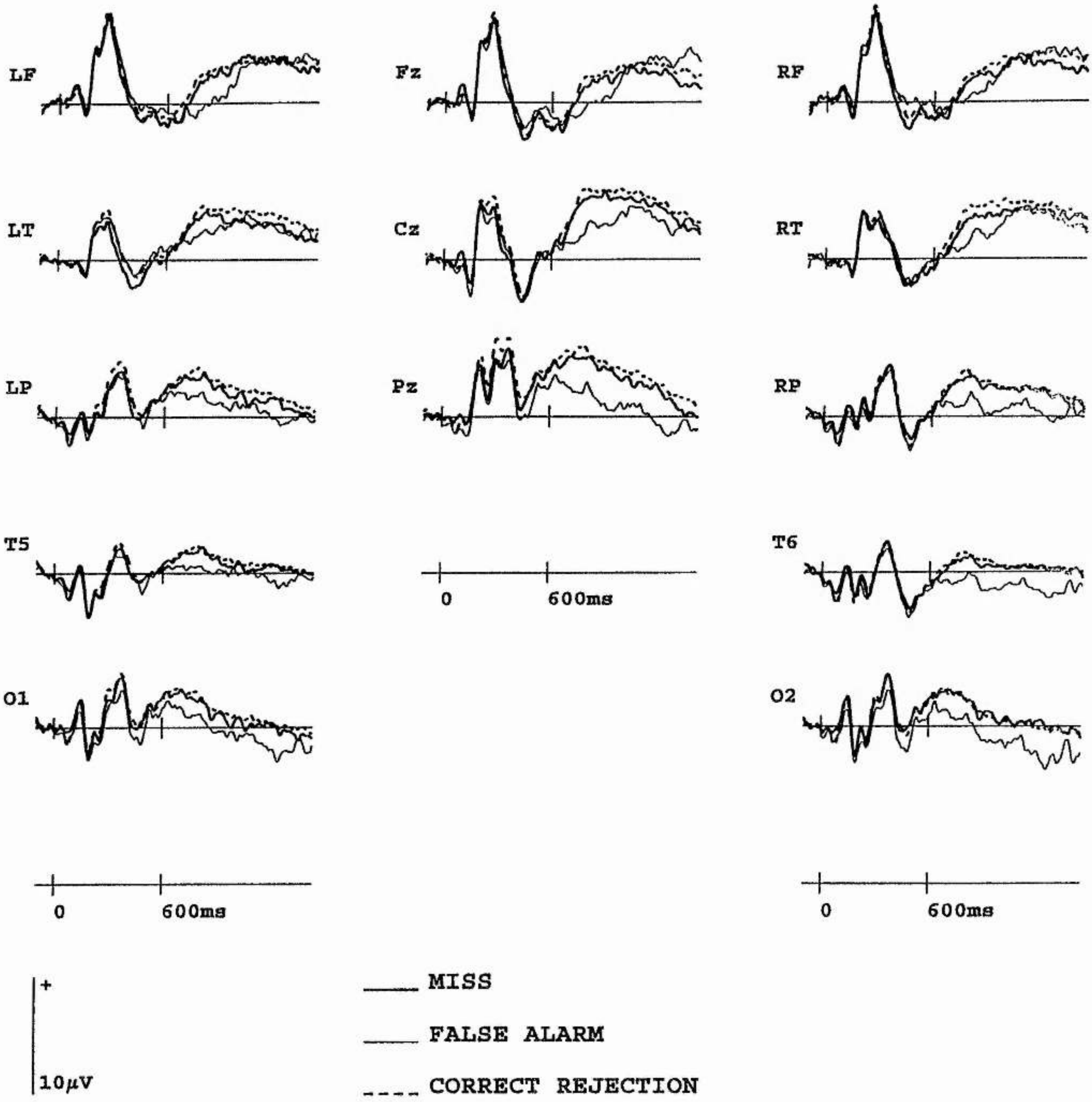


Figure 8.3 Grand average ERPs associated with the miss, false alarm, and correct rejection response categories in experiment 4. Electrode sites as for figure 5.1.

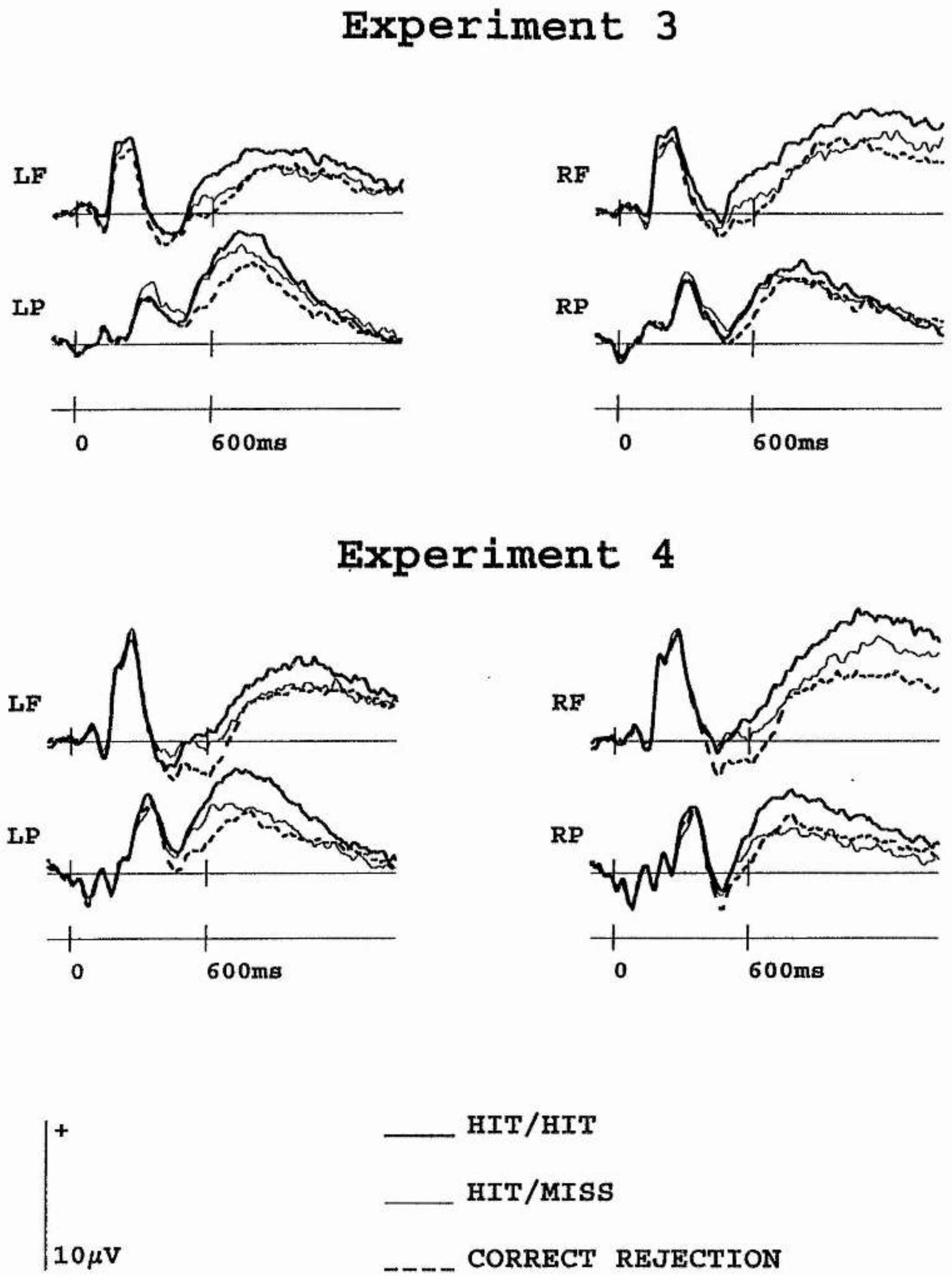


Figure 8.4 Grand average ERPs associated with the hit/hit, hit/miss, and correct rejection response categories in experiments 3 and 4. Electrode sites displayed are left and right frontal and parietal.

Chapter 9

9 *An ERP study of the processes supporting judgements on a recognition memory exclusion task*

9.1 *Introduction*

In this experiment ERPs were recorded whilst subjects performed a recognition memory exclusion task. The format of this task was introduced in chapter 1 during the discussion of the process-dissociation procedure (PDP), and will be reviewed briefly here. In common with the task used in experiments 1-4, the exclusion task requires subjects to discriminate between new test items and between old items which have been presented in one of two contexts in a previous study phase. However, unlike the task employed in experiments 1-4, these discriminations are concatenated in a single forced-choice binary test judgement. This is achieved by requiring subjects only to respond 'old' to items presented in one of the two study contexts (targets), and to respond 'new' to words spoken in the alternate context (non-targets), as well as to genuinely new words.

When the PDP was introduced in chapter 1, a description of the assumptions underlying the putative bases for judgements to targets and non-targets in an exclusion task was given. The description will be expanded here, since the results of experiments 3 and 4 permit predictions to be made regarding the ERP old/new effects that should be observed in a recognition memory exclusion task.

According to Jacoby and colleagues (Jacoby, 1991; Jacoby *et al.*, 1993), correct judgements to targets (target hits) can be made either on the basis of recollection, or of fluency, whilst correct judgements to non-targets (non-target hits) can be made on the basis of recollection, or if the test word is forgotten, since both of these cases will result in a correct new judgement to a non-target. Given the results of experiments 3 and 4, it would therefore be predicted that both parietal and frontal old/new effects should be revealed by a comparison of the ERPs to correct rejections with the ERPs to target and non-target hits.

Further predictions can also be made regarding the ERP old/new effects for incorrect judgements to targets and non-targets. Incorrect judgements to targets (target misses) should, according to the assumptions underlying the PDP, only be made if test items are forgotten. Given that experiments 1 to 4 have revealed no reliable differences between the ERPs to misses and to correct rejections, it is predicted that no differences should be revealed by the comparison of these response categories in this experiment.

Finally, incorrect judgements to non-targets (non-target false alarms) are assumed to be made on the basis of fluency (Jacoby, 1991). It was previously noted (chapter 1) that since responses of this type can be viewed as correct old judgements which are unaccompanied by retrieval of accurate contextual information, these responses can in principle be made on the basis of the type of recognition unaccompanied by retrieval of context which is proposed by proponents of the declarative memory view (e.g. Squire, 1982a). However, regardless of whether the basis for these responses is assumed to be fluency or this latter type of recognition, the old/new effects for this response category should be similar to those reported for the hit/miss response categories in experiments 3 and 4

9.2 *Methods*

Subjects: A total of 27 subjects participated in the experiment, for which they were paid £3.00/hour. The data from 1 subject was discarded because the subject failed to follow the test instructions. The data from a further two subjects were excluded due to excessive EOG artifact and an unstable electrode site respectively. Of the remaining 24 subjects, 12 were female. 23 were right-handed, as defined by writing hand. The age of subjects ranged from 17 to 27 years (average age 20).

Experimental Material: Stimuli consisted of 360 words. The procedure for forming study and test lists was the same as for experiment 3. The sole difference was that no non-words were employed at study, each study list therefore consisting of 180 words. The study lists were split into 3 blocks of 60 items, and within each block half of the words were spoken in the male voice, and half were spoken in the female voice. The structure of test lists was identical to that of experiment 3. The characteristics of the visually and auditorily presented stimuli were as for experiment 3.

Procedure: At study subjects heard one of the four study lists. In the subsequent test phase subjects saw one of the two test lists which corresponded to the list they had encountered at study.

Each subject performed two interleaved study tasks. In task A subjects judged whether the auditorily presented word was pleasant or unpleasant, whilst in task B subjects judged whether each word was active or passive. They were informed that there were no right or wrong answers. This task manipulation was introduced in order to improve memory for the context in which words had been encountered at study.

Each trial was preceded by one of two cues, an 'O' or an 'X'. The former denoted that the subject should make a task A judgement to the subsequent word, whilst the latter indicated that the subject should make a task B judgement. The fixation cue was removed 100 msec prior to stimulus onset. Following presentation of the word subjects were required to respond verbally with the voice of the speaker (male/female) and the task judgement (active/passive or pleasant/unpleasant). The mapping between voice and task was constant for each subject and was counterbalanced across subjects. Following the verbal response, subjects pressed a key to initiate the next trial. Subjects were asked to remain relatively still and to fixate on the screen for the duration of the task, but no instructions to restrict eye blinks to a portion of the trial were given.

At test subjects made a single binary judgement by depressing one of two microswitches on which their index fingers rested. An asterisk preceded presentation of each stimulus and was removed 100 msec prior to stimulus onset. Each subject was instructed only to respond old to words spoken in the male/female voice at study (targets), and to respond new to words spoken in the alternate voice (non-targets), as well as to genuinely new words. The voice designated as the target was counterbalanced across subjects. The fixation point reappeared on the screen 2 seconds after the response of the subject.

Subjects were instructed to make the old/new judgement as quickly and as accurately as possible. The hands used for the judgements were counterbalanced across subjects, so that there was no correlation between hand and response type.

EEG Recording: All aspects of the recording procedure and criteria for EEG and EOG were as for experiment 3, as were all amplifier characteristics.

9.3 *Results*

9.31 *Behavioural data*

Table 9.1 displays the probability of a correct response for new words, targets, and non-targets. Two measures of discrimination were calculated in order to assess whether subjects could reliably distinguish between the three classes of test word. The first discrimination measure was calculated by subtracting the probability of an incorrect judgement to a non-target from the probability of a correct target judgement. The second discrimination measure followed the same procedure, with the exception that the probability of an incorrect judgement to a non-target was replaced by the probability of a false alarm. The analyses revealed that subjects were able reliably to distinguish between targets and non-targets ($t(23) = 8.01$; $p < .001$), and between targets and new words ($t(23) = 11.28$; $p < .001$).

A preliminary analysis revealed that memory for words spoken in either the male or the female voice was equivalent, consistent with the findings of experiment 3. It was therefore possible to derive estimates of recollection and fluency from this exclusion task, as described by Yonelinas and Jacoby (1994). Using the inclusion and exclusion equations given in chapter 1, and assuming independence, the derived estimates of recollection and fluency were: $R = 0.32$, and $F = 0.41$.

Table 9.2 displays the RTs for correct and incorrect test judgements to the three classes of test word (targets, non-targets, and new words). ANOVA on the RTs employed the factors of response accuracy (correct vs incorrect) and word type (target vs non-target vs new). The analysis revealed main effects of both factors (respectively, $F(1,23) = 40.14$; $p < .001$, and $F(1.6,36.4) = 5.24$; $p < .05$), and an interaction between these factors ($F(2.0,44.9) = 19.76$; $p < .001$).

Post-hoc analyses of the interaction term revealed that correct judgements to new words were faster than correct judgements to old words (targets and non-targets), whilst for incorrect judgements RTs to targets were reliably faster than RTs to non-targets and RTs to new words. Further, correct responses were faster than incorrect responses for new words and non-targets, but not for targets.

Analyses of the variability of the reaction time distributions also revealed an effect of accuracy ($F(1,23) = 5.65$; $p < .05$), and an interaction between accuracy and word type ($F(1.9,44.0) = 4.88$; $p < .05$). *Post-hoc* analyses revealed that whereas there was less variability in the RTs for correct rejections than for false alarms, the RT variability for old words did not differ according to response accuracy. Further, whilst the RT variability for incorrect judgements did not differ according to word type, for correct judgements the RT variability for correct rejections was reliably smaller than that for non-target hits.

9.32 *ERP Analyses*

The ERPs to the critical response categories are collapsed across study voice, on the basis of the findings in experiments 3 and 4 that the ERPs did not differ according to this factor. The initial analyses reported below consist of a comparison of the non-target hit, target miss, and correct rejection ERPs. Following the analysis strategy employed in experiment 1, these ERPs were initially compared in a global ANOVA including the three response categories. Any reliable effects involving response category were followed up by subsidiary ANOVAs comparing these ERPs on a pairwise basis. The analyses were performed over the same time windows employed in experiments 3 and 4.

Of the 24 subjects who took part in the experiment, only 20 made sufficient misses to permit formation of reliable averaged ERPs for this response category. Consequently, the analyses described below for the correct rejection, target miss, and non-target hit ERPs are for these 20 subjects. Figure 9.1 displays the ERPs to these three response categories. At left hemisphere sites the non-target hit ERPs are more positive than the ERPs to correct rejections from approximately 400-900 msec post-stimulus. At right hemisphere sites over these latencies the correct rejection and non-target hit ERPs differ little. From 900 msec onwards these ERPs differ little at left hemisphere sites, whilst over the right hemisphere the ERPs to correct rejections are more positive-going. The differences between the ERPs to correct rejections and target misses are negligible for the duration of the recording epoch.

The global ANOVAs over the 500-800 and 800-1100 msec time windows revealed no effects involving response category at the midline. At lateral sites, the analyses revealed interactions between category and hemisphere (500-800: $F(1.6,30.1) = 4.50$; $p < .05$; 800-1100: $F(1.4,27.1) = 7.00$; $p < .01$). Over the 1100-1400 msec epoch the analyses revealed a main effect of response category at midline ($F(1.6,30.0) = 6.06$; $p < .01$) and lateral sites ($F(1.5,29.4) = 3.99$; $p < .05$). The midline analysis also revealed an interaction between category and site ($F(2.5,48.4) = 4.75$; $p < .01$).

The results of the three sets of paired analyses on the non-target hit, target miss and correct rejection ERPs are displayed in table 9.3, which tabulates all effects involving response category. The analyses were performed over the same 3 time windows employed in experiments 3 and 4. Note that, as in previous experiments, although table 9.3 displays the results of all analyses at midline and at lateral sites, the text below refers only to the analyses over those latency regions and those locations where the global analyses revealed reliable effects involving response category. For these three classes of ERPs appendix 1.5 displays mean amplitude measures for the three critical response categories for each electrode site over the three latency regions analysed.

9. 321 *Analysis of non-target hit, target miss, and correct rejection ERPs*

Comparison of the ERPs to target misses and those to correct rejections revealed no reliable effects involving response category over any of the latency regions analysed. Over the 500-800 msec epoch the comparisons of these ERPs to the ERPs to non-target hits revealed interactions between response category and hemisphere. In both cases the interaction reflects the fact that the non-target hit ERPs are more positive over the left hemisphere, whilst at right hemisphere sites there is little differentiation between the ERPs to these three response categories.

Consistent with these findings, a planned comparison of the non-target hit and correct rejection ERPs at the left and right parietal sites revealed a response category x site interaction ($F(1,19) = 4.52$; $p < .05$), reflecting the fact that the differences between these ERPs are larger at the left (1.3 μV) than at the right (0.1 μV) hemisphere site.

Over the 800-1100 msec epoch interactions between response category and hemisphere were again revealed by the comparisons of the non-target hit and target miss ERPs, and the non-target hit and correct rejection ERPs. Over this epoch the interaction terms reflect the fact that at left hemisphere sites the ERPs to these three response categories differ little, whilst at right hemisphere sites the non-target hit ERPs are more negative-going (see appendix 1.5).

Over the 1100-1400 msec epoch the comparison of the ERPs to non-target hits and to misses revealed an interaction between category and site at the midline, reflecting the fact that while these ERPs differ little at Fz, the non-target hit ERPs are more negative-going at Cz and Pz. The analysis of these ERPs at lateral sites revealed a main effect of category, reflecting the fact that the ERPs to non-target hits are more negative-going.

Finally, over the 1100-1400 msec epoch the comparison of the non-target hit and correct rejection ERPs at lateral sites revealed a main effect of response category, and an interaction between this factor and hemisphere. The interaction reflects the fact that whilst the non-target hit ERPs are more negative over both hemispheres, they are relatively more negative at right hemisphere sites. At the midline, the analysis of these ERPs revealed a main effect of category, and an interaction between this factor and site. A Scheffé analysis revealed that the size of the differences between these ERPs is larger at Cz and Pz than at Fz, with the non-target hit ERPs being more negative in both cases. A planned comparison of the non-target hit and correct rejection ERPs at the left and right frontal electrode sites revealed no effects involving response category.

9.322 *Analysis of old/new effects for words correctly assigned to study context*

One of the predictions made in the introduction to this experiment was that, on the basis of the hit/hit old/new effects observed in experiments 3 and 4, both the target hit and non-target hit old/new effects should consist of a parietal and a frontal component. This prediction was assessed directly by an analysis of the old/new effects for these ERPs at the left parietal site over the 500-800 msec epoch, and the right frontal site over the 1100-1400 msec epoch. These electrode locations were selected on the basis of the findings in experiments 3 and 4 that they were the most sensitive indices, over the 500-800 and 1100-1400 msec epochs respectively, of the parietal and frontal old/new effects.

This analysis was performed on the data from 23 subjects who contributed a minimum of 16 artifact free trials to each of these response categories.

The ERPs to the target hit, non-target hit and correct rejection response categories are displayed in figure 9.2. The pattern of differences between the non-target hit and correct rejection ERPs is very similar to that described for figure 9.1 above. The target hit ERPs are more positive than the ERPs to correct rejections from 300-400 msec post-stimulus. The differences between these ERPs are characterised by the left greater than right parietal asymmetry evident in the comparison of the non-target hit and correct rejection ERPs. In addition, in comparison to the ERPs to correct rejections the target hit ERPs are also characterised by the right-greater left frontal positivity reported for the hit/hit ERPs in experiments 3 and 4.

ANOVA comparing the target hit, non-target hit, and correct rejection ERPs at the left parietal site over the 500-800 msec epoch revealed a main effect of response category ($F(2.0,37.6) = 11.98$; $p < .001$). *Post-hoc* analyses (Newman Keuls) revealed that the non-target hit ERPs were reliably more positive than the ERPs to correct rejections, and that the target hit ERPs were more positive than both of these response categories.

The analysis at the right frontal site over the 1100-1400 msec epoch also revealed a main effect of response category ($F(1.9,35.7) = 10.69$; $p < .001$). In this case *post-hoc* analyses (Newman Keuls) revealed that whereas the ERPs to correct rejections and non-target hits did not differ, both were less positive than the target hit ERPs. The mean amplitudes for these three response categories are shown in figure 9.3, which displays the mean amplitudes at the left parietal site over the 500-800 msec epoch, and those at the right frontal site over the 1100-1400 msec epoch.

9.323 *Analysis of ERPs to false alarms and non-target false alarms*

Of the 24 subjects who took part in the experiment, only 9 made sufficient false alarms to permit formation of reliable averaged waveforms, and only 13 subjects made sufficient non-target false alarms, therefore precluding a direct comparison of these ERPs. The ERPs to non-target false alarms and to correct rejections are displayed in appendix 2.4, in order to permit a visual comparison of these effects with the hit/miss old/new effects reported in experiments 3 and 4.

The ERPs to the two false alarm response categories were collapsed together in order to permit a comparison of these ERPs to those to correct rejections. As for previous analyses the collapsed category was formed by computing a weighted average of the false alarm and non-target false alarm ERPs for each subject. The ERPs to these response categories are shown in figure 9.4. The ERPs to the collapsed false alarms are more negative-going from approximately 650 msec post-stimulus, and the magnitude of this negativity is larger at right than at left hemisphere locations.

Following the analysis procedure outlined above, the correct rejection and false alarm ERPs were compared at the left parietal site over the 500-800 msec epoch, and the right frontal site over the 1100-1400 msec epoch. Analyses of variance directly comparing these ERPs at these sites revealed no effects involving response category. [4]

9.4 *Discussion*

The pattern of ERP findings in the exclusion task is similar to that reported in experiments 3 and 4, where a somewhat different task manipulation was employed. The qualitative similarities between the findings in the two experiments therefore suggest that the previous findings were not solely a consequence of the particular task in which ERPs to recollected and unrecollected words were recorded. In particular, the frontal old/new effect cannot be directly related to the fact that in experiments 3 and 4 the old/new judgement preceded the context judgement: qualitatively similar frontal old/new effects are evident when the old/new and context judgements are combined in a single binary decision.

The comparison of the non-target hit, target miss and correct rejection ERPs revealed that the latter two response categories were not reliably different, whereas the non-target hit ERPs showed a reliable parietal old/new effect. Given that the responses for these three categories were made on the same key, these effects cannot be due to different response requirements to old and new words. Further, for the same reason the results are inconsistent with the view that the parietal old/new effect is a consequence of a

[4] In this experiment the only ERP effects on which it was possible to perform topographic analyses were the changes over time in the differences between the ERPs associated with the target hit and the correct rejection response categories. These analyses are not reported here as they provide no new information to the results of the topographic analyses on words correctly assigned to study context which were reported in experiments 3 and 4.

mismatch between the probability of old and new responses to recognition memory test items. Early proposals for the functional significance of the parietal ERP old/new effect linked it to the properties of the P300 (Karis *et al.*, 1984; Neville *et al.*, 1986). By this view, the parietal old/new effect represents a larger P300 to correctly recognised old items. The effect was proposed to arise because old items were associated with a lower response probability than were new items, and were more likely to be regarded as targets than were new words. These two factors have been linked to the amplitude of the P300 component (Pritchard, 1981).

To address these proposals, Smith and Guster (1993) systematically varied the probability of the occurrence of old words and reported no reliable differences between the magnitude of the old/new effects observed when the ratio of old to new items was 4:1 or 1:4. In the task employed by Smith and Guster (1993), subjects responded only to targets, which in different experimental conditions were either designated as new or old words. Comparison of the ERPs when either new or old words were designated as targets revealed old/new effects of equivalent magnitude. These findings are hard to reconcile with the view that the old/new effect is solely a modulation of the P300 component.

The findings in this experiment extend those of Smith and Guster (1993), who recorded ERPs from three midline electrode sites only (Fz, Cz, Pz). In the present study ERPs were recorded from sites over both hemispheres. Given that previous reports of the parietal old/new effect have reported the left-greater-than-right asymmetry at parietal sites (Neville *et al.*, 1986; Rugg and Doyle, 1992), it is important to demonstrate that the lateralisation of the effect is still evident when response probability is controlled for.

9.41 *Comparison of target and non-target hits*

The second set of analyses performed on the ERPs in the exclusion task specifically investigated the parietal and frontal old/new effects for two classes of ERPs: target hits and non-target hits. The analyses of these ERPs revealed that both response categories were associated with a parietal old/new effect, whilst only the target hit response category was associated with a reliable frontal old/new effect.

There are a number of possible explanations for the differences between the old/new effects for these two response categories. First, the differences may be related to the proportion of trials contributing to each category which were not associated with recollection. It is possible to estimate these proportions using the value of R which was calculated from the PDP equations above. The PDP estimated the probability of recollection to be 0.32, whilst the probabilities of correct responses to targets and non-targets were 0.58 and 0.74 respectively. Therefore, if the estimate provided by the PDP equations is reliable, then slightly more than 50% of the trials comprising the target hit response category were associated with recollection, whilst slightly less than 50% of the trials comprising the non-target hit response category were associated with recollection.

9.42 *Non-target hit old/new effects*

For the non-target hits, the trials which were not associated with recollection should be words which were forgotten, since either recollecting or forgetting a non-target should lead to a correct 'new' response. It seems unlikely that the impact of the ERPs to forgotten words would be to influence the morphology of the differences between the ERPs to non-target hits and correct rejections, since the ERPs to target misses were not reliably different to those for correct rejections over any recording epoch. If this interpretation is correct, then the old/new effects for the non-target ERPs reflect the distribution of neural activity which differentiates recollected non-targets and correctly recognised new words.

9.43 *Target hit old/new effects*

Trials contributing to the ERPs to target hits can either be based upon recollection, or on the basis of a guess when subjects are in possession of sufficient information to make a correct old judgement, but insufficient information to make a correct context judgement. This latter response type is directly comparable to the hit/miss response category introduced previously.

In experiment 4, where the introduction of the don't know response option was intended to make a clear distinction between recollected and unrecollected words, the hit/miss ERPs were associated with parietal and frontal old/new effects which were reliably smaller than the hit/hit ERP old/new effects, although the scalp distributions of the old/new effects for these response categories were not reliably different. If this relation between recollected and unrecollected words also holds for the trials comprising the target hit response category, then it is reasonable to assume that the differences between the target hit and correct rejection ERPs are reflective of the distribution of neural activity that differentiates recollected targets and correctly recognised new words.

However, a further factor which may contribute to the ERP old/new effects for the target hit ERPs is the mismatch in the probability of making an old or a new judgement to test words. Whilst the preceding sections converge on the conclusion that the parietal old/new effect is not solely related to the functional properties of the P300, it is still possible to argue that the frontal old/new effect is sensitive to the asymmetry between the probabilities of making an old or a new judgement to test stimuli on an exclusion task. In the experiment only 25% of test items were targets, and the probability of a 'new' response was 0.72.

There are two lines of evidence which oppose the view that the frontal old/new effect for the target hit ERPs is due to a response probability confound. First, the comparison of the ERPs to correct rejections and the ERPs to the response category formed by collapsing the ERPs to false alarms and non-target false alarms revealed no reliable differences between these ERPs at the right frontal site. The ERPs comprising this collapsed response category are all associated with a response made on the same key as the ERPs to target hits. Therefore, if the frontal old/new effect were solely related to response probability, there should be no evidence that responses made on this key are associated with different experimental effects.

A second line of argument is the lack of equivalence between the scalp distributions of the frontal old/new effect and the P300. Whilst the former is maximal at frontal sites and displays a right-greater-than-left asymmetry, the peak amplitude of the latter is typically at centro-parietal scalp locations (Pritchard, 1981). The differences in scalp distribution therefore converge with the findings of the comparison of false alarms and correct rejections to suggest that the differences between the ERPs to correct rejections and target hits are an accurate reflection of the neural activity differentiating recollected targets and correctly recognised new words.

9.44 *The relationship between frontal and parietal old/new effects*

The preceding observations suggest that the differences between the frontal old/new effects for the target and non-target hit response categories may not be related to differences in the morphology of single trial ERPs contributing to the averaged ERPs, or to a confound between the proportions of stimuli which are to be designated old or new in the exclusion task. The differences between the old/new effects for these response categories therefore shed light on the relationship between the parietal and frontal old/new effects observed in experiments 3 to 5.

In experiments 3 and 4 the parietal and frontal old/new effects were larger for the hit/hit than for the hit/miss ERPs. These findings are consistent with the view that the size of these topographically distinct old/new effects covaries. However, the disproportionate attenuation of the frontal old/new effects revealed by the comparison of target and non-target hit ERPs suggests that the relationship between these effects is not necessarily of this form. The findings suggest that the magnitude of the frontal old/new effect depends upon factors related to the task in which context discriminations are made.

The experimental findings are consistent with the view that recollected words are processed differently depending upon whether they have been designated pre-experimentally as targets or non-targets. It is not clear what form this differential processing would take. However, one possibility is that subjects set different decision criteria for targets and non-targets. For example, subjects may have required recovery of more information to make an old judgement to a target than that required to make a 'new' judgement to a non-target. This interpretation is consistent with the proposal advanced in the last chapter that the frontal old/new effect is sensitive to retrieval of information in a graded rather than an all-or-none fashion, with larger frontal old/new effects associated with retrieval of more information. A related possibility is that the processing indexed by the frontal old/new effect is not solely related to memory retrieval *per se*. Rather, the effect may be sensitive to manipulations of a more strategic nature. The implications of this proposal will be returned to in the following chapter (Chapter 10), and the general discussion (Chapter 11).

9.45 *The process-dissociation procedure*

The differences between the old/new effects for the target hit and non-target hit ERPs are also relevant to the process-dissociation procedure, introduced by Jacoby and colleagues as a means of estimating the contributions of recollection and fluency to performance on recognition memory tasks (Jacoby, 1991; Jacoby *et al.*, 1993). As has been previously noted, the PDP applies mathematical descriptions to the processes which are assumed to contribute to memory judgements. The computation of the relative contributions of recollection and fluency depends upon a comparison of performance under inclusion and exclusion instructions. This comparison can be made either on the basis of performance on a single exclusion task (Yonelinas, 1994; Yonelinas and Jacoby, 1994), or in separate inclusion and exclusion tasks (Jacoby, 1991).

In order for reliable estimates of recollection and fluency to be obtained, the PDP must assume that the processes of recollection and fluency are engaged to the same extent under inclusion and exclusion instructions (Jacoby, 1991; Toth *et al.*, 1995). The ERP data from this experiment are consistent with the view that subjects process recollected words differently according to their designation as targets or non-targets. If this is the case then the assumption of the equivalence of recollection is questionable when inclusion and exclusion scores are obtained from a single exclusion task. If the differences between the ERP old/new effects for the target and non-target hits index differential processing that results in different behavioural responses to targets and non-targets, then the PDP will yield estimates which are not representative of the processes contributing to task performance.

The experimental findings do not speak directly to the question of whether the assumption of equivalence of recollection is valid when separate inclusion and exclusion tasks are performed, although as noted in chapter 1, this assumption has been questioned on theoretical grounds (Graf and Komatsu, 1994). It would be of considerable interest to record ERPs, in the same subjects, under inclusion and exclusion instructions, in order to assess directly the processing afforded recollected items in these two conditions.

In principle, the process-dissociation procedure offers a means of estimating the relative contributions of the processes which contribute to recognition memory judgements. This is an important endeavour, since a more complete understanding of the functional attributes of these processes, and the relationship between them, is only possible if there is some reliable means of assessing their contributions to performance on different tasks and in different subject populations. It remains to be seen whether the recent criticisms of the PDP will help to formulate a modified framework from which reliable and consistent estimates can be obtained.

Table 9.1 Probability of correct judgements to targets, non-targets and new words in experiment 5. (s.d. in brackets)

	Word Type		
	New	Target	Non-Target
P(Correct)	0.85(0.12)	0.58(0.15)	0.74(0.13)

Table 9.2 Reaction times (msec) for correct and incorrect judgements to targets, non-targets and new words in experiment 5.

	Response	Word Type		
		New	Target	Non-Target
RT	Correct	1140	1357	1348
SD		404	464	517
RT	Incorrect	1538	1285	1490
SD		510	507	488

Table 9.3 Results of the pairwise analyses of the non-target hit, target miss, and correct rejection ERPs in experiment 5. The analyses were performed over the 500-800, 800-1100, and 1100-1400 msec epochs.

	500-800 msec				800-1100 msec				1100-1400 msec			
Non-Target Hit vs CR **Midline**	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category (1,19)	3.73	6.45		0.069	6.73	6.07		**0.018**	24.37	4.14		**0.001**
Category x Site (2,38)	0.64	0.86	0.61	n.s.	2.85	1.17	0.59	0.091	8.63	1.47	0.68	**0.004**
Lateral												
Category (1,19)	2.99	9.61		0.100	1.91	10.16		n.s.	14.00	5.89		**0.001**
Category x Site (4,76)	0.65	1.40	0.51	n.s.	0.71	1.51	0.50	n.s.	1.04	1.74	0.54	n.s.
Category x Hem (1,19)	7.05	2.45		**0.016**	16.58	1.94		**0.001**	7.12	2.21		**0.015**
Category x Hem x Site (4,76)	1.38	0.63	0.55	n.s.	2.32	0.71	0.55	n.s.	1.37	0.93	0.53	n.s.
Target Miss vs CR	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Midline												
Category	0.00	7.28		n.s.	2.28	6.13		n.s.	1.17	11.07		n.s.
Category x Site	2.04	0.95	0.65	n.s.	0.75	1.20	0.85	n.s.	1.04	1.19	0.87	n.s.
Lateral												
Category	0.39	11.83		n.s.	0.31	11.40		n.s.	0.03	16.85		n.s.
Category x Site	0.96	1.70	0.34	n.s.	0.42	1.81	0.36	n.s.	0.33	2.10	0.37	n.s.
Category x Hem	0.11	1.86		n.s.	0.53	1.67		n.s.	0.32	3.16		n.s.
Category x Hem x Site	1.08	0.47	0.75	n.s.	0.72	0.78	0.67	n.s.	0.52	1.22	0.68	n.s.
Non-Target Hit vs Target Miss **Midline**	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	2.28	11.19		n.s.	0.76	9.21		n.s.	3.99	10.45		0.061
Category x Site	0.83	1.72	0.54	n.s.	1.68	1.40	0.62	n.s.	4.05	1.86	0.69	**0.043**
Lateral												
Category	0.59	17.39		n.s.	0.35	18.26		n.s.	4.49	15.51		**0.048**
Category x Site	0.15	2.45	0.42	n.s.	0.04	2.04	0.53	n.s.	0.89	2.55	0.48	n.s.
Category x Hem	4.94	4.30		**0.039.**	5.19	4.32		**0.035**	1.57	5.58		n.s.
Category x Hem x Site	1.54	0.64	0.62	n.s.	2.36	0.79	0.67	0.089	1.11	1.23	0.70	n.s.

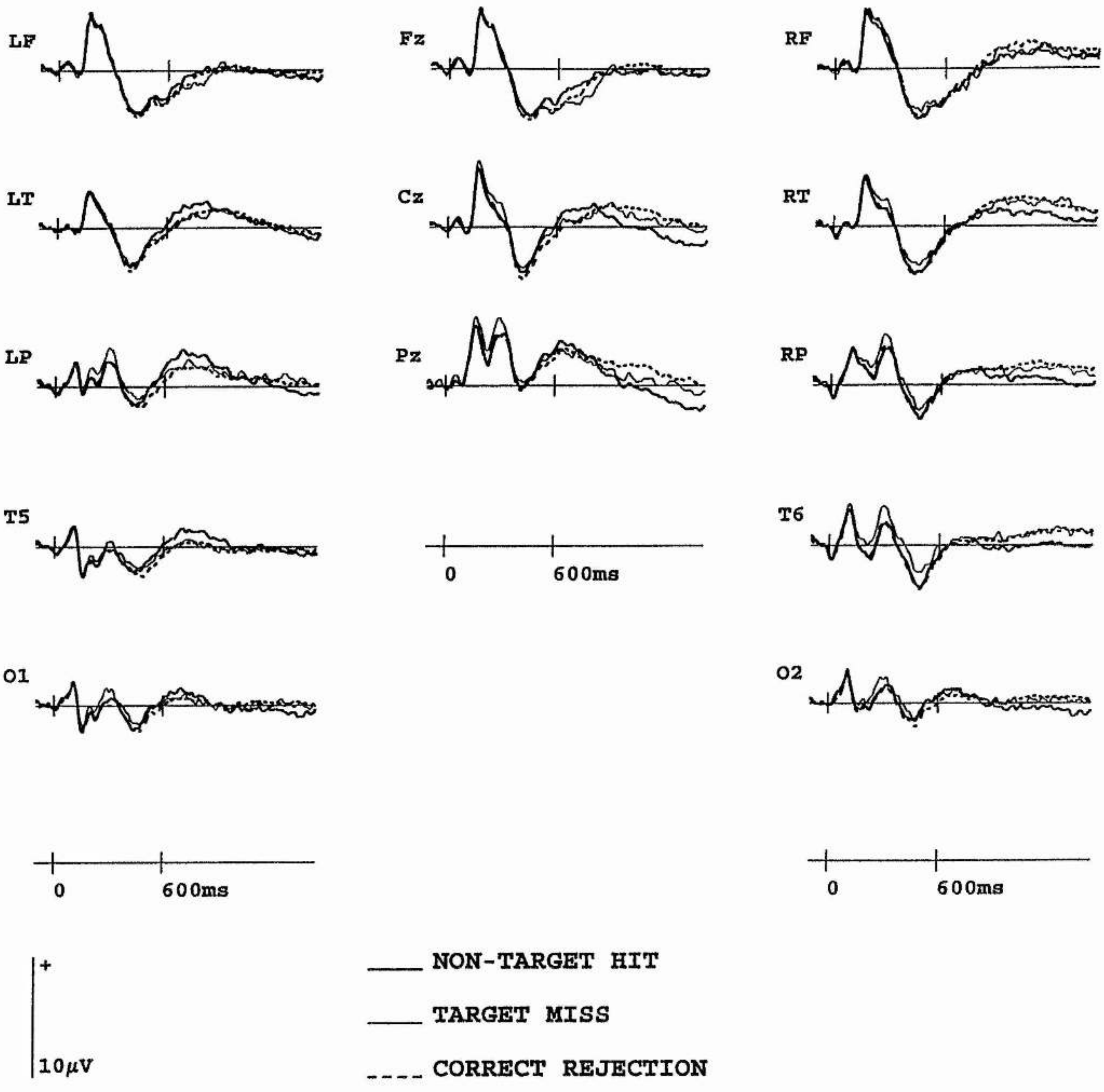


Figure 9.1 Grand average ERPs associated with the non-target hit, target miss, and correct rejection response categories in experiment 5. Electrode sites as for figure 5.1.

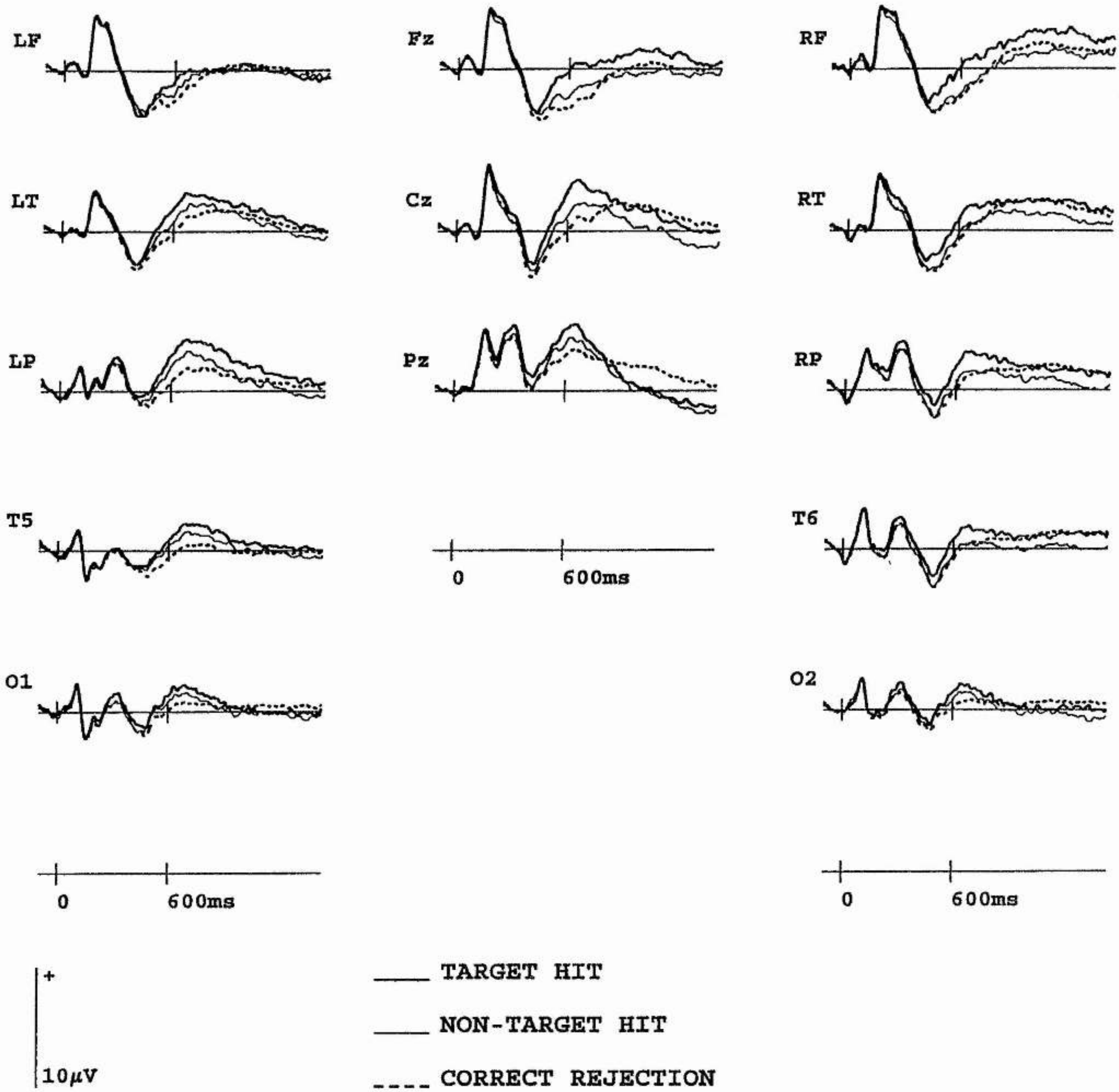
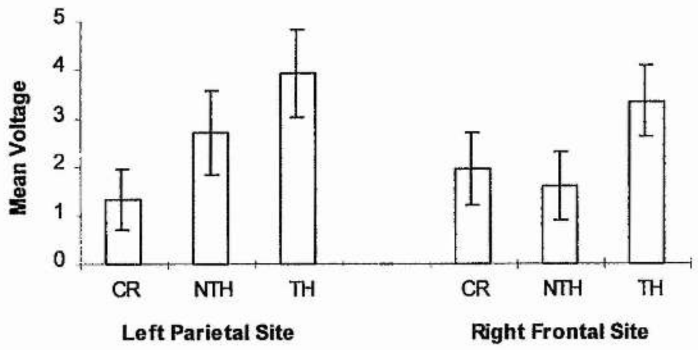


Figure 9.2 Grand average ERPs associated with the target hit, non-target hit, and correct rejection response categories in experiment 5. Electrode sites as for figure 5.1.

Figure 9.3 Mean amplitudes of the target hit (TH), non-target hit (NTH), and correct rejection (CR) ERPs in experiment 5. The figure displays the mean amplitudes of these ERPs at the left parietal site over the 500-800 msec epoch, and the right frontal site over the 1100-1400 msec epoch.



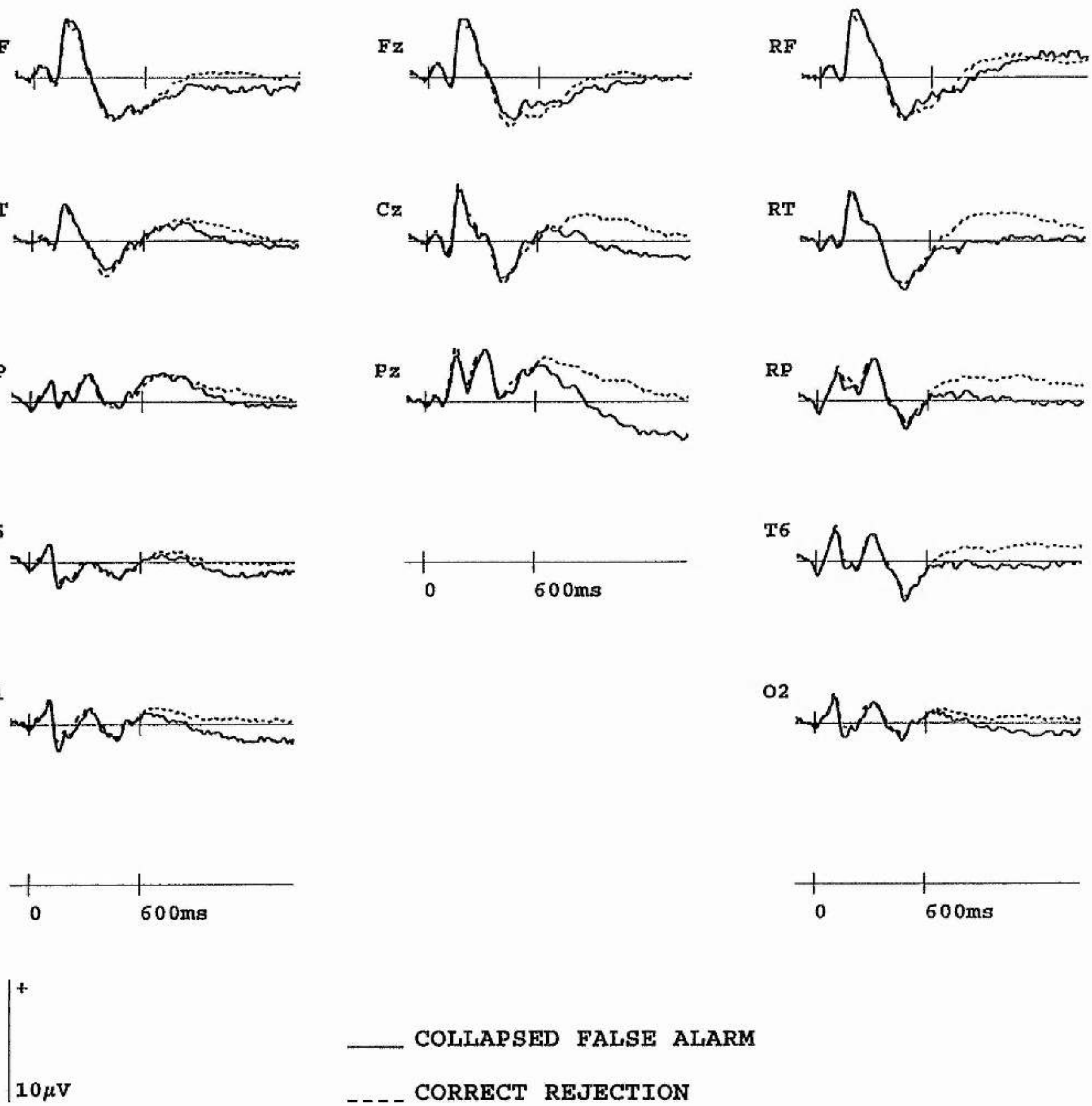


Figure 9.4 Grand average ERPs for the correct rejection response category, and the response category formed by collapsing across the false alarm and non-target false alarm response categories in experiment 5. Electrode sites as for figure 5.1.

Chapter 10

10 *An ERP study of the processes supporting discriminations between spoken and heard words*

10.1 *Introduction*

The results of experiments 3 and 4 are consistent with the view that two distinct processes contribute to memory for study context. Experiment 5 revealed similar effects in a different experimental paradigm, and in addition suggested that the parietal and frontal old/new effects can vary independently. In these three experiments the same contextual manipulation - speaker voice - was employed. Consequently, the extent to which the parietal and frontal old/new effects are specific to retrieval of voice information has not been directly addressed, and the experiments do not speak directly to the question of the sensitivity of the parietal and frontal old/new effects to different types of contextual information.

These issues were investigated in this experiment, where a different contextual manipulation was introduced. The experimental procedure was similar to that of experiment 3, with the exception that the male/female voice discrimination was replaced by a contextual manipulation in which subjects discriminated between words that they had heard at study, and words that they had spoken aloud.

The requirement to distinguish between spoken and heard words has been employed in what are termed the *reality monitoring* and *source monitoring* frameworks (Johnson and Raye, 1981; Johnson, 1993). Reality monitoring refers to discriminations between memories for internally generated events and memories for externally derived information. Reality monitoring is subsumed within the general framework of source monitoring, which refers to the processes which contribute to judgements about the origin of memories (Johnson, Hashtroudi and Lindsay, 1993). The framework assumes that context or source judgements are made on the basis of an assessment of a number of qualitatively different forms of retrieved information about a prior episode. These forms of retrieved information can be employed for source discriminations since they are assumed to be differently distributed in memories of different origins. For example, the memory trace for a heard word will contain more perceptual information than a memory trace for a word that a subject imagined hearing. Conversely, the imagined event will contain more details of the cognitive operations performed during the act of imagining (Johnson, Foley and Leach, 1988).

The aim of this experiment was to investigate the sensitivity of ERPs to the different forms of contextual information that are represented in memories for heard and spoken words. In a parallel to the example given above comparing the contextual information available for source judgements for perceived and imagined events, it is reasonable to suppose that whilst both spoken and heard words will be associated with contextual information regarding speaker voice, words spoken at study will also be associated with information regarding the cognitive operations engaged during the production and generation of a verbal response.

As previously noted, the results of experiments 3 and 4 are consistent with the view that there is a direct relationship between the size of the parietal and frontal old/new effects, and the findings of experiment 5 suggest that the size of the frontal old/new effect can vary independently of the parietal effect. Comparing hit/hit ERPs to words which were either spoken or heard at study is another means of investigating the relationship between the parietal and frontal old/new effects, since source judgements to spoken and heard words may be made on the basis of, or be accompanied by, retrieval of different forms of information.

10.2 *Methods*

Subjects: 17 subjects took part in the experiment, for which each was paid £3.50/hour. One subject was discarded from the final analyses due to excessive EOG artifact and head movement. Of the remaining 16 subjects, 8 were female. All subjects were right-handed. The age range of subjects was 18 to 34 years (average age 24).

Experimental Materials: Stimuli consisted of 300 words. The characteristics of visually and auditorily presented stimuli were as for experiment 4. The 300 words were divided into 10 equal lists, each of which formed one study list. Within each study list, half of the words were to be spoken by the subject, whilst the other half were to be heard.

Test lists were formed by combining two word lists. Each test list was paired with one study list such that each test list consisted of 30 studied words, and 30 words which would be presented at test for the first time. Two random orderings were applied to each test list, such that each study list mapped onto two test lists. In addition, a further 10 study lists were formed by reversing those words which were to be heard and those words which were to be spoken. Therefore each word was spoken on one study list and heard on a second study list. Two filler words were presented at the beginning and the end of each study list, and a further two fillers were presented at the start of each test list.

Task lists (consisting of 5 study and 5 test lists) were formed from the pool of 20 study and 20 test lists. The list for each subject was constructed such that, across the 5 study-test cycles, all words presented within a task list were presented on only one study list and one test list. The experimental design was balanced for item and order effects, and partially balanced for order of block presentation. Subjects encountered a total of 170 items at study (30 critical items and 4 fillers per study list). At test subjects encountered 310 items (60 critical items and 2 fillers per test list).

Procedure: Prior to the study phase subjects were informed that they were to take part in a task consisting of 5 study-test blocks. They were informed that in each study phase they would see words presented one at a time, and would subsequently be required either to listen to the word or to pronounce the word aloud. Subjects were informed that in a following test phase they would be required to distinguish between old and new words, and between words they had spoken and words they had heard at study.

Each study phase trial commenced with the presentation of a fixation point on the TV monitor (either an 'O' or an 'X'), which also served as the cue indicating whether subjects were to pronounce the word to be presented visually on the screen or to listen to the word as it was repeated. An 'O' indicated that the word should be spoken, whilst an 'X' indicated that the word would be heard through the headphones. The fixation point was removed from the screen 100 msec prior to stimulus presentation, and the visual stimulus was exposed for 300 msec. This was replaced 1 second later by a question mark, which served as the cue for the subject either to pronounce the visually presented word, or to listen as the word was spoken through the headphones. The study task was experimenter paced, in order to control for variation in the amount of time taken by subjects to make a verbal response. Subjects were asked to remain relatively still during each study phase, but they were not instructed to restrict their eye blinks to any portion of each study trial.

In each test block the experimental design was identical to that of experiment 3, with the exception that the context judgement in this case was a discrimination between words that had been spoken at study and words that had been heard.

EEG Recording: EEG and EOG recording and criteria were as for experiment 4, with the exception that EEG was recorded from 25 scalp sites. This montage consisted of those sites employed in experiment 4 (17 electrode sites), and an additional 8 sites: left and right superior frontal and temporal (F3, F4, C3, C4), and left and right inferior frontal and temporal (F7, F8, T3, T4).

10.3 *Results*

In the following sections, words correctly judged old to which a correct spoken/heard judgement was made will be referred to as belonging to the *spoken hit/hit* and *heard hit/hit* response categories respectively. The corresponding terms for incorrect source judgements are *spoken hit/miss* and *heard hit/miss*. Words which the subject pronounced aloud at study will be referred to as *spoken* words, and those which the subject listened to will be referred to as *heard* words.

10.31 *Behavioural data*

Table 10.1 displays the probability of a correct response for new and old test words. The old words are separated as a function of study task (spoken/heard). For both classes of old word, discrimination was above chance level (for spoken words $t(15) = 25.59$; $p < .001$; for heard words $t(15) = 16.67$; $p < .001$). Comparison of the discrimination measures for these classes of old words revealed a discrimination advantage for words which subjects repeated aloud ($t(15) = 7.52$; $p < .001$).

ANOVA on the RTs for old and new items (table 10.2) employed the factors of response accuracy (correct vs. incorrect) and word status (spoken vs heard vs new). The ANOVA revealed a main effect of accuracy ($F(1,15) = 22.58$; $p < .001$), and an interaction between word type and accuracy ($F(1.4,21.7) = 6.87$; $p < .01$). *Post-hoc* tests (Newman Keuls) revealed that whereas there were no reliable differences between the RTs to words associated with correct judgements, for incorrect judgements words heard at study were responded to more quickly than were new words. Further, whilst there were no reliable differences between the RTs to correct and incorrect judgements for heard words, for spoken words and for new words the RTs to correct judgements were reliably faster than the RTs to incorrect judgements.

Table 10.3 displays the conditional probability of a correct source judgement for words correctly judged old. Also displayed (far right column) is the probability of a 'spoken' source judgement to a false alarm. The probability of a correct source judgement was 0.80, a value reliably above the chance probability of 0.50 ($t(15) = 11.13$; $p < .001$). The probability of a 'spoken' judgement to a new word incorrectly judged old was significantly less than 0.50 ($t(15) = 2.68$; $p < .001$).

Table 10.4 displays the RTs to words correctly judged old, separated according to the accuracy of the subsequent source judgement. Too few incorrect source judgements were made to permit a comparison of these RTs separated according to study manipulation (spoken/heard). A direct comparison of the RTs for correct source judgements, separated according to study manipulation, revealed that words spoken at study which were correctly recognised and correctly assigned to source were associated with faster RTs than words which were heard ($F(1,15) = 17.84$; $p < .01$).

10.32 *ERP Analyses*

The ERP old/new effects associated with the spoken hit/hit and heard hit/hit response categories are displayed in figure 10.1. This figure displays the ERPs for the 13 sites comprising the standard montage. Appendix 2.6 displays the ERPs for these three response categories for the 25 electrode sites from which EEG was recorded. Appendix 2.5 displays hit/hit and hit/miss old/new effects for 9 subjects who made sufficient incorrect source judgements to permit formation of reliable averaged ERPs for these response categories. The hit/hit and hit/miss ERPs are collapsed across study context. As only 9 subjects contributed to these grand average ERPs no analyses were performed. However, as appendix 2.5 illustrates, the old/new effects for these response categories are qualitatively similar to those reported in experiments 3 to 5.

Figure 10.1 shows that both the spoken and heard hit/hit response categories are characterised by a parietal and a frontal old/new effect. Whilst the frontal old/new effects are of equal magnitude for these two response categories, the parietal old/new effect is larger for the spoken hit/hit ERPs.

These old/new effects were analysed by three paired comparisons, the results of which are shown in table 10.5. The analyses were performed over the same three epochs employed in experiments 3 to 5. Note that whilst table 10.5 displays the results of the analyses at midline and at lateral sites, the text below refers only to the analyses at lateral sites, except where the midline analyses contribute additional information from that which is revealed by the analyses at lateral scalp locations. The mean amplitudes of these ERPs over the three epochs are displayed in appendix 1.6.

10.321 *Spoken hit/hit old/new effects*

Over the 500-800 msec epoch the analysis of the spoken hit/hit and correct rejection ERPs revealed a main effect of response category, reflecting the fact that the spoken hit/hit ERPs are more positive. The analyses also revealed response category x site and response category x hemisphere interactions. Scheffé analyses revealed that the differences between these ERPs are reliably larger over the left hemisphere than over the right (4.2 μV vs 2.7 μV), and that the differences between these ERPs are larger at parietal sites than at anterior temporal and occipital sites.

A planned comparison of these ERPs at the left- and right-parietal sites revealed a main effect of category, and an interaction between category and site (respectively $F(1,15) = 41.95$; $p < .001$, and $F(1,15) = 13.28$; $p < .01$), reflecting the markedly asymmetric (left > right) old/new effect for the spoken hit/hit ERPs.

800-1100 msec: The analyses over this epoch revealed a main effect of response category, and an interaction between category and site, reflecting the fact that the hit/hit ERPs are more positive, with the largest differences between these ERPs occurring at the frontal electrode sites.

1100-1400 msec: Over this epoch the comparisons of these ERPs revealed an interaction between response category and site, reflecting the fact that at posterior sites the hit/hit ERPs are more negative than those to correct rejections, whilst these ERPs are more positive at frontal sites. A planned analysis of the differences between these ERPs at the left- and right- frontal sites revealed no evidence for the right-greater-than-left frontal asymmetry reported previously (experiment 4).

10.322 H*eard hit/hit old/new effects*

Over the 500-800 msec epoch the analyses revealed a main effect of response category, reflecting the fact that the hit/hit ERPs are more positive. A planned analysis of the old/new effects at the left- and right- parietal sites revealed a main effect of response category ($F(1,15) = 9.10$; $p < .01$), and an interaction between this factor and site ($F(1,15) = 5.70$; $p < .05$), reflecting the fact that these old/new effects are larger at the left parietal site than at it's contralateral homologue.

The analyses over the 800-1100 and 1100-1400 msec epochs both revealed interactions between response category and site. Over the former epoch the interaction term reflects the fact that the hit/hit ERPs are more positive at frontal locations, but differ little from the ERPs to correct rejections at more posterior scalp sites. Over the latter epoch the hit/hit ERPs remain more positive than the ERPs to correct rejections at frontal locations, but are markedly more negative than the ERPs to correct rejections at posterior scalp sites. As for the spoken hit/hit old/new effects, a planned analysis of the differences between the heard hit/hit and correct rejection ERPs at the left- and right-frontal sites revealed no evidence for the right-greater-than-left frontal asymmetry reported previously.

10.323 *Comparison of spoken and heard hit/hit old/new effects*

The only significant differences involving response category revealed by the comparison of the spoken and heard hit/hit ERPs were over the 500-800 msec epoch. The analyses revealed a main effect of response category, and an marginally non-significant interaction ($p = 0.05$) between this factor and hemisphere. These results reflect the fact that the spoken hit/hit ERPs are more positive, with the difference between these ERPs tending to be larger over the left hemisphere than over the right. Consistent with these results, a planned comparison of these ERPs at the left and right parietal sites revealed a main effect of response category ($F(1,15) = 18.06$; $p < .01$), and an interaction between category and site ($F(1,15) = 5.97$; $p < .05$), reflecting the fact that the differences between these ERPs are larger at left than at right parietal locations.

10.324 *Topographic analyses*

A single ANOVA was employed to compare the scalp distributions of the spoken hit/hit and heard hit/hit old/new effects across the 500-800 and 1100-1400 msec epochs. The analysis revealed no effects involving the factor of category, but gave rise to a significant interaction between epoch and site ($F(3.2,47.8) = 9.93$; $p < .001$), reflecting a change in scalp distribution over time. An analysis of these ERPs restricted to the sites comprising the standard montage also revealed an interaction between epoch and site ($F(3.5,53.1) = 7.44$; $p < .001$).

As for the previous 3 experiments, this interaction was further investigated by a restricted analysis on the rescaled data for the frontal and parietal sites over the two epochs. The subsidiary analysis revealed an interaction between epoch and site (F(1,15)

= 11.36; $p < .01$), and the interaction between epoch and hemisphere approached significance ($F(1,15) = 4.25$; $p = .06$). These interactions reflect the anterior shift in the distribution of the old/new effects over time, which is more marked at right than at left hemisphere sites.

The temporal evolution of the old/new effects for the spoken and heard hit/hit ERPs can be seen clearly in the topographic scalp maps displayed in figures 10.2 and 10.3. These maps are depictions of the distribution of the old/new effects obtained by subtracting the ERPs to correct rejections from the ERPs to the spoken hit/hit and heard hit/hit ERPs respectively. The maps portray the mean distribution of voltage across the scalp for the latency regions indicated. Note that each map displays relative voltage for the particular latency region - the voltage range for the maps is displayed alongside the grayscale on the right hand side of each map. The voltage values which lie between the 25 marked electrode sites were derived using a spherical spline interpolation (Perrin, Pernier, Bertrand and Echallier, 1989; Perrin, Pernier, Bertrand, Giard and Echallier, 1987).

Figure 10.2 displays the scalp distribution of the spoken and heard hit/hit old/new effects across the 500-800 and 1100-1400 msec epochs. It can be seen that the distributions of these effects differ little, consistent with the finding that the topographic analysis revealed no effects involving response category. Figure 10.3 displays the distribution of the old/new effects for the collapsed spoken and heard hit/hit ERPs. The maps display the distribution of these effects for this collapsed response category across 4 epochs: 500-700, 700-900, 900-1100, and 1100-1400 msec. The shift in the distribution over time from a left parietal to a right-frontal maximum is clearly evident.

10.4 *Discussion*

10.41 *Behavioural results*

The rationale for employing the blocked experimental design was to engender high levels of recognition memory and memory for study context. The high hit rate and low false alarm rates indicate that the design achieved this aim. For the initial old/new judgement, discrimination was greater for words which had been spoken at study. In contrast, for the context judgement words which had been heard were associated with a higher probability of a correct context judgement. In addition, analysis of the context judgements to false alarms revealed a large response bias towards responding 'heard'. This finding is consistent with previous studies which have found that under conditions of uncertainty subjects are more likely to judge that an item or event was associated with an external source than an internal one - a finding which has been termed the 'it-had-to-be-you' phenomenon (Johnson *et al.*, 1993; Johnson and Raye, 1981; Johnson, Raye, Foley and Foley, 1981). Given the evidence for the large response bias revealed here it is unclear whether the context memory advantage for words which were heard at study is due to genuinely better memory for those items, or alternatively is due to a predisposition to make a 'heard' response when uncertain of the study context for an item that has been correctly judged old (see Batchelder and Riefer, 1990, and comments in chapter 5 regarding response biases of this form). For the purposes of the following discussions, the critical import of this response bias is the likely impact on the averaged heard hit/hit ERP waveforms. This will be returned to below, following a summary of the principal ERP findings.

10.42 *ERP results*

The principal experimental comparison was between the old/new effects for the spoken hit/hit and heard hit/hit ERPs. Topographic analyses of these ERPs revealed no reliable differences, suggesting that the same processes contribute to correct context judgements for these two response categories. The topographic analyses did reveal evidence for changes in scalp distribution over time when the differences between the hit/hit and correct rejection ERPs were compared over the 500-800 and 1100-1400 msec epochs. These changes in topography with time are depicted in figures 10.2 and 10.3.

The principal difference between the spoken hit/hit and heard hit/hit ERPs was the size of the parietal old/new effect. The analyses of these ERPs at lateral sites revealed that the spoken hit/hit ERPs were reliably more positive than the heard hit/hit ERPs at parietal locations, with the magnitude of the difference being larger over the left hemisphere than over the right. The larger parietal old/new effect for the spoken hit/hit ERPs contrasts with the absence of any reliable differences between these ERPs from approximately 900 msec onwards, although over these later epochs both classes of hit/hit ERPs were reliably more positive than the ERPs to correct rejections.

One possible explanation for the attenuated parietal old/new effect for the heard hit/hit response category is that, in comparison to the spoken hit/hit ERPs, a larger proportion of the trials comprising this category were not associated with retrieval of veridical contextual information. This interpretation is supported by the finding that the probability of a 'heard' context judgement to a false alarm was markedly larger than the probability of a 'spoken' judgement, which, as noted above, suggests that when subjects were uncertain of the study context to which to assign a correctly recognised old word, they were more likely to respond 'heard' than 'spoken'. By this view, the attenuated parietal old/new effect for the heard hit/hit ERPs reflects the influence of correct guesses on the resulting averaged ERPs.

In order for this explanation to hold, it is necessary to assume that the parietal old/new effect, but not the frontal old/new effect, is attenuated when subjects make a correct recognition judgement in the absence of veridical contextual information. This is the case since the frontal old/new effect is statistically equivalent for the spoken hit/hit and heard hit/hit response categories. This interpretation is inconsistent with the findings in experiments 3 and 4 that correctly recognised words which are not accompanied by veridical contextual information are associated with smaller parietal *and* frontal old/new effects than are correctly recognised words which are also correctly assigned to study context. Therefore, if the differences between the spoken and heard hit/hit ERPs were related to the proportion of correct guesses comprising the heard hit/hit ERPs, a diminution of parietal and frontal old/new effects would be predicted.

If the differences between these two classes of ERPs are not due to any systematic variation in the proportion of single trials of different morphology contributing to the averaged hit/hit ERPs, than these findings support the view, advanced on the basis of the results of experiment 5, that the relationship between the magnitude of the parietal and frontal old/new effects is task dependent. The results of the two experiments represent a double dissociation: in experiment 5 a disproportionate attenuation of frontal old/new effects relative to differences in parietal old/new effects was observed, whereas in this experiment differences in magnitude were evident only for the parietal old/new effect.

One explanation for the differences between the two hit/hit old/new effects at parietal locations is that they reflect differences in the amount or type of information retrieved from memory. The view that the parietal old/new effect is sensitive to the amount of information retrieved has previously been proposed (chapters 7 and 8) on the basis of the findings that the parietal old/new effect for hit/hit ERPs is larger than that for hit/miss ERPs.

In order for this interpretation to hold for the spoken and heard hit/hit old/new effects it is necessary to assume that pronouncing a word aloud at study elicits retrieval of more information in a subsequent test phase than does simply hearing a word at study. This appears to be a plausible assumption, since only the former involves the production of a verbal response. The additional retrieved information for spoken words would presumably include information regarding response production, or information related to the cognitive operations undertaken in order to generate the verbal response (Johnson *et al.*, 1988).

If this interpretation of the spoken and heard parietal hit/hit old/new effects is correct, then the question becomes why the frontal old/new effects for these two response categories are not reliably different. The parietal and frontal old/new effects have previously been discussed in terms of their relationship to the processes of retrieval and integration/elaboration which have been proposed to be necessary for retrieval of contextual or source information (Moscovitch, 1992; Moscovitch, 1994; Squire, 1992). In the particular version of this model proposed by Moscovitch (1992, 1994), the first of these processes - the retrieval function - is reflexive, suggesting that the information retrieved is not dictated by task demands. By contrast, the second process, although involving integration of information delivered by the retrieval function, is under strategic (conscious) control, and should therefore be modulated by the demands of particular tasks.

The differences between the parietal and frontal old/new effects for spoken and heard words can be accommodated within this framework. First, the larger parietal old/new effects for spoken words reflects retrieval of more, or qualitatively different, information, as discussed above. Second, the frontal old/new effects reflect the processing of the retrieved information which is *necessary* for the task discrimination. By this view, whilst words spoken at study are associated with retrieval of more information at test, only a proportion of that information is employed for the task judgements. This may reflect the fact that the processing indexed by the frontal old/new effect acts selectively over a proportion of the retrieved information which is relevant to the task judgement. Alternatively, a proportion of the information retrieved for words spoken at study may not be relevant to the task distinction. Whilst the ERP data cannot distinguish between these possibilities, the central point is that both assume the frontal old/new effect is more sensitive to task demands than the parietal old/new effect.

Table 10.1 Probabilities of correct old/new judgements for test words in experiment 6. Old words are separated according to study manipulation (s.d. in brackets).

	Word Type		
	Spoken	Heard	New
P(Correct Judgement)	0.88(0.12)	0.73(0.16)	0.91(0.08)

Table 10.2 Reaction times (msec) for initial old/new judgements to new and old words in experiment 6. Old words are separated according to study manipulation.

	Response	Word Type		
		Spoken	Heard	New
RT	Correct	1161	1235	1140
SD		253	252	273
RT	Incorrect	1332	1235	1459
SD		218	235	232

Table 10.3 Probabilities of correct source judgements for words correctly judged old in experiment 6. Also displayed (far right column) is the probability of a 'spoken' judgement to a false alarm (s.d. in brackets).

	Study Manipulation		False Alarms
	Spoken	Heard	P(Spoken Judgement)
P(Correct)	0.74(0.13)	0.86(0.14)	0.18(0.26)

Table 10.4 Reaction times (msec) for words correctly judged old in experiment 6. The RTs are separated according to the accuracy of the subsequent source judgement.

	Response	Study Manipulation	
		Spoken	Heard
RT	Correct	1129	1229
SD		258	270
RT	Incorrect	1265	1170
SD		201	124

Table 10.5 Results of the pairwise analyses of the spoken hit/hit, heard hit/hit, and correct rejection ERPs in experiment 6. The analyses were performed over the 500-800, 800-1100, and 1100-1400 msec epochs.

	500-800 msec				800-1100 msec				1100-1400 msec			
Spoken Hit/Hit vs CR **Midline**	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category (1,15)	39.84	12.78		**0.001**	0.42	22.82		n.s.	4.80	24.71		**0.045**
Category x Site (2,30)	0.29	2.58	0.55	n.s.	7.31	3.63	0.74	**0.007**	15.39	3.45	0.67	**0.001**
Lateral												
Category (1,15)	57.50	16.48		**0.001**	5.37	19.77		**0.035**	2.03	10.28		n.s.
Category x Site (4,60)	4.22	2.15	0.44	**0.030**	4.51	2.66	0.51	**0.019**	7.38	2.89	0.60	**0.001**
Category x Hem (1,15)	6.97	6.76		**0.019**	1.69	12.24		n.s.	0.49	11.69		n.s.
Category x Hem x Site (4,60)	2.92	1.00	0.48	0.072	1.49	2.10	0.50	n.s.	2.43	1.46	0.51	n.s.
Heard Hit/Hit vs CR **Midline**	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	11.02	16.15		**0.005**	0.65	23.79		n.s.	0.79	38.55		n.s.
Category x Site	0.07	1.79	0.85	n.s.	9.53	2.21	0.98	**0.001**	21.03	1.69	0.95	**0.001**
Lateral												
Category	10.32	25.22		**0.006**	3.65	20.51		0.076	0.12	28.98		n.s.
Category x Site	1.52	1.64	0.50	n.s.	5.06	1.91	0.54	**0.011**	10.14	2.39	0.59	**0.001**
Category x Hem	3.24	4.13		0.092	1.02	8.74		n.s.	1.11	7.06		n.s.
Category x Hem x Site	2.31	0.65	0.51	n.s.	0.99	1.73	0.45	n.s.	1.62	1.21	0.56	n.s.
Spoken Hit/Hit vs Heard Hit/Hit **Midline**	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*	*F*	*MSE*	ε	*p*
Category	11.09	7.68		**0.005**	0.10	6.42		n.s.	3.40	8.54		0.085
Category x Site	0.42	1.56	0.66	n.s.	0.20	2.20	0.76	n.s.	0.91	1.95	0.73	n.s.
Lateral												
Category	13.20	16.24		**0.002**	0.22	12.49		n.s.	0.42	17.27		n.s.
Category x Site	1.94	1.16	0.51	n.s.	0.43	1.45	0.40	n.s.	0.40	1.66	0.46	n.s.
Response category x Hem	4.53	2.26		0.050	0.98	2.49		n.s.	0.05	2.80		n.s.
Category x Hem x Site	0.89	0.53	0.53	n.s.	1.27	0.57	0.73	n.s.	1.70	0.56	0.69	n.s.

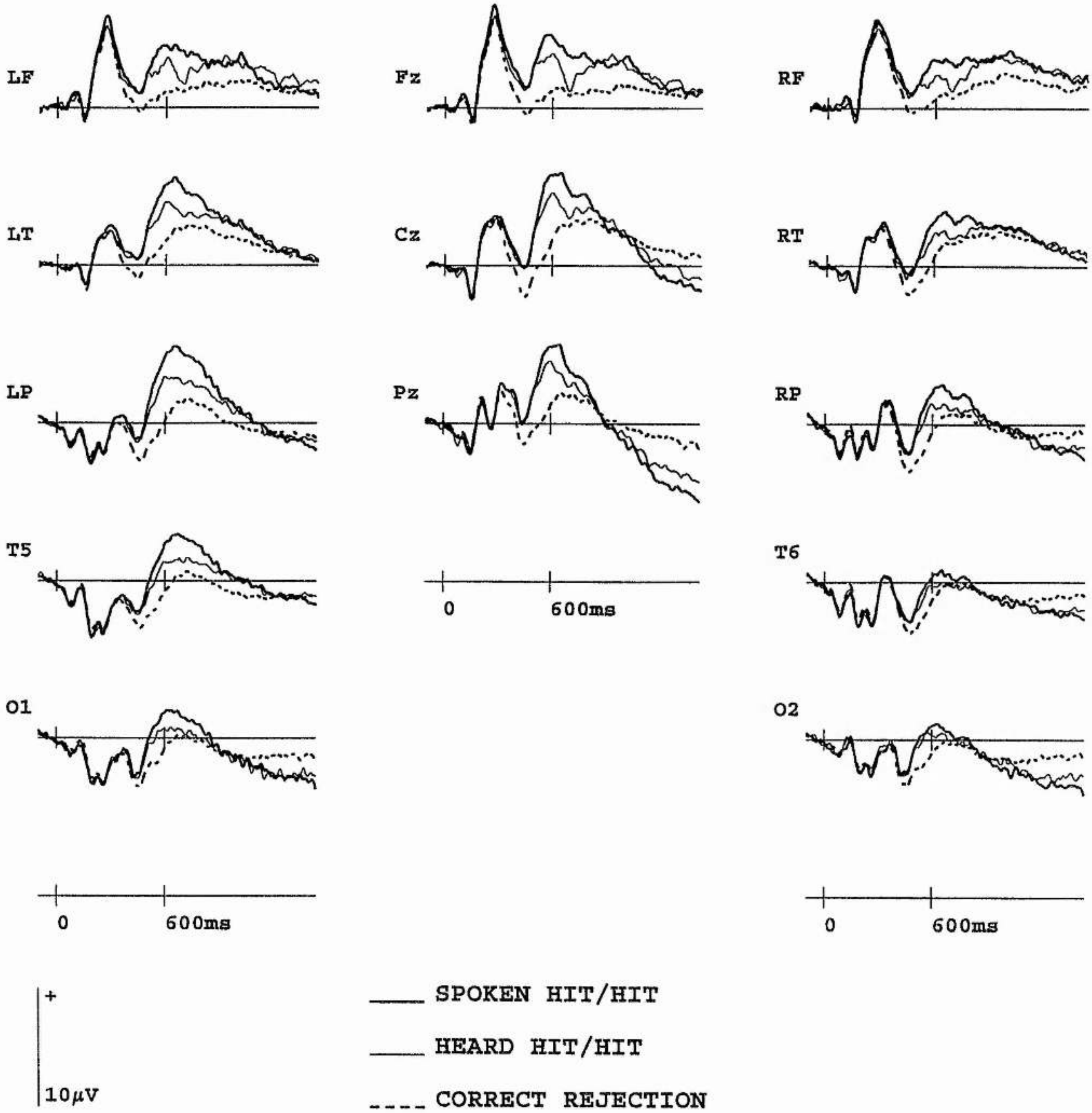


Figure 10.1 Grand average ERPs associated with the spoken hit/hit, heard hit/hit, and correct rejection response categories in experiment 6. Electrode sites as for figure 5.1.

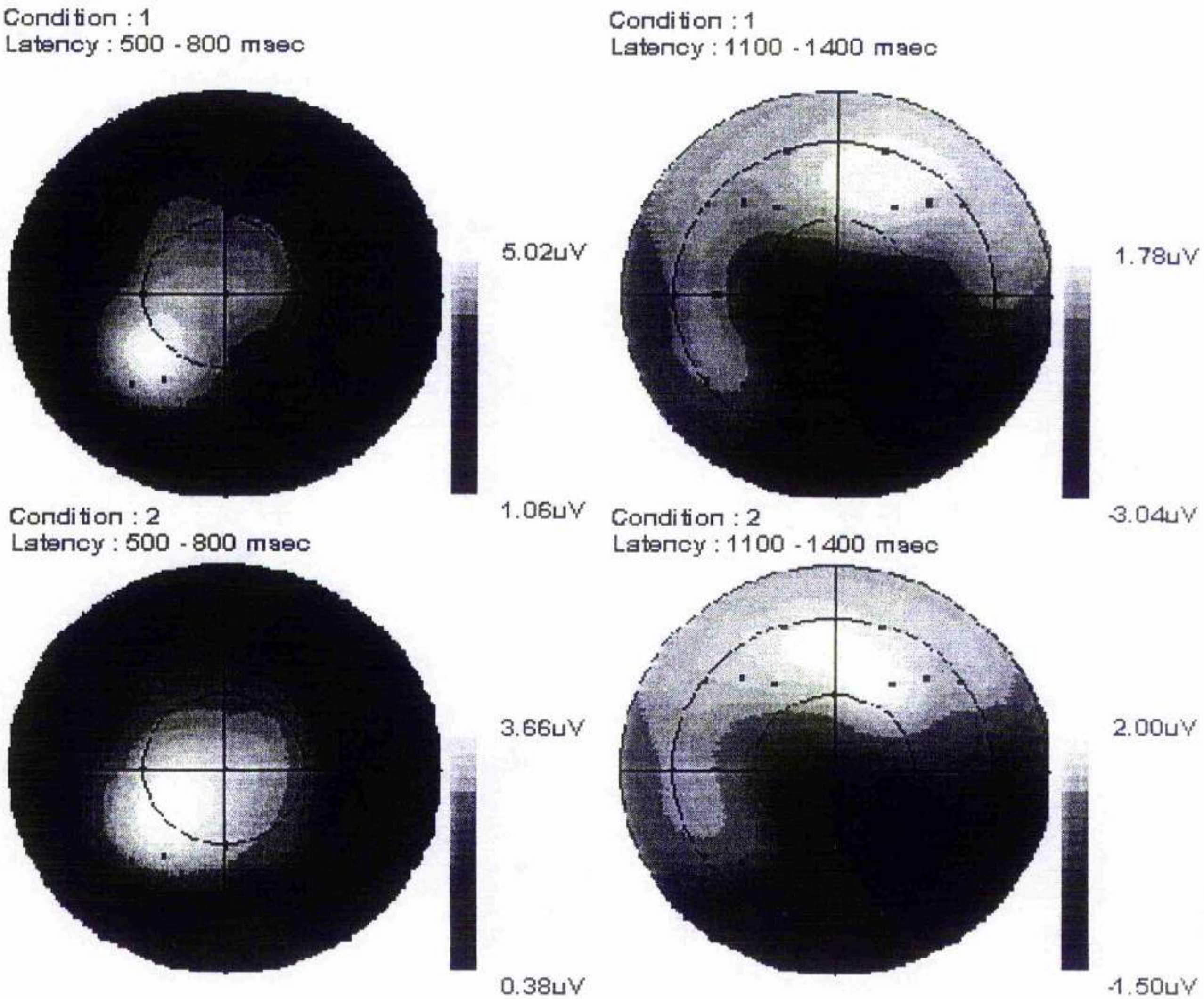


Figure 10.2 Topographic maps showing the relative amplitudes of the differences between the ERPs to correct rejections and the spoken (condition 1) and heard (condition 2) hit/hit ERPs. The maps display the relative amplitudes over two latency regions: 500-800 and 1100-1400 msec.

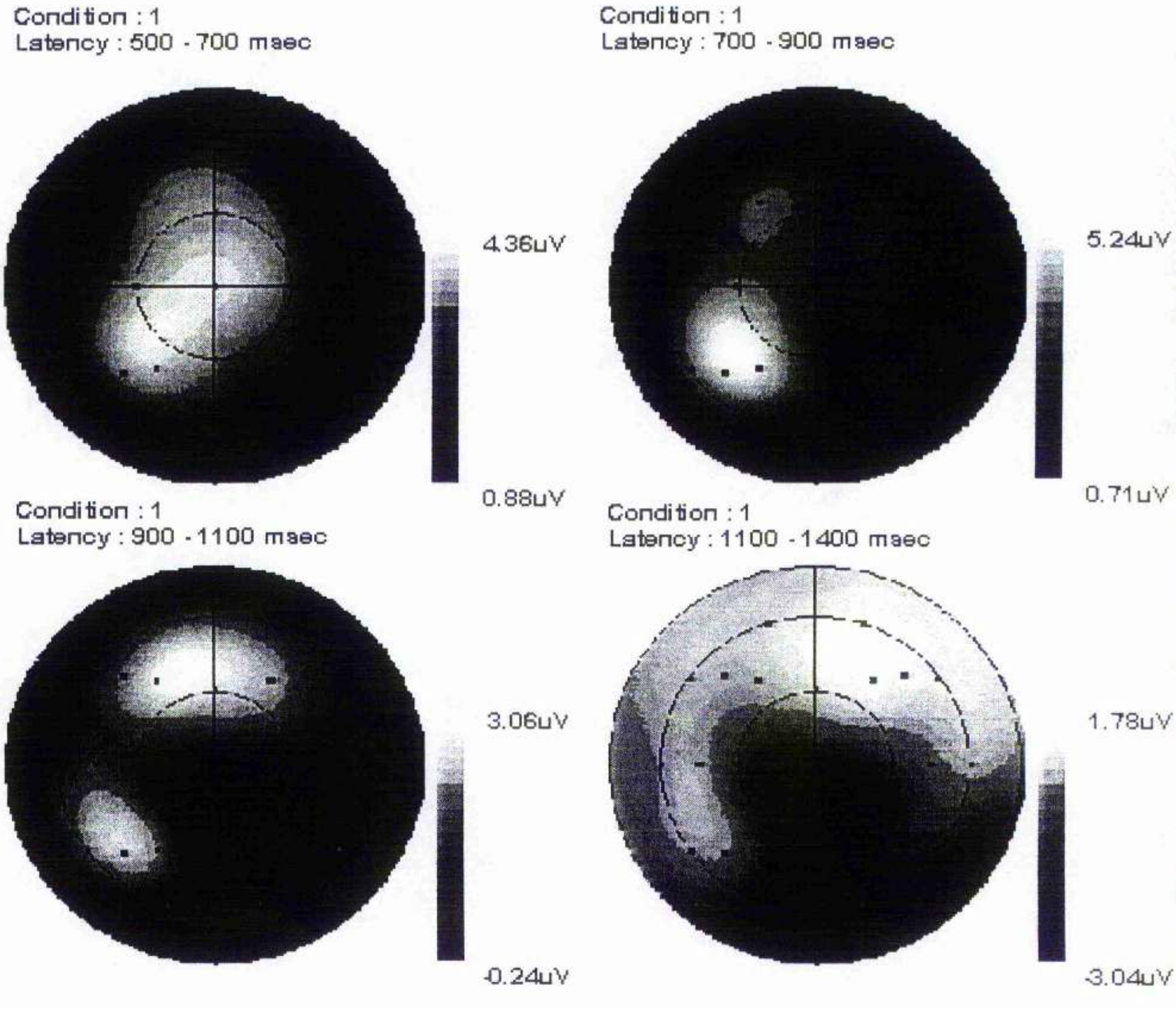


Figure 10.3 Topographic maps showing the relative amplitudes of the differences between the collapsed hit/hit and correct rejection ERPs. The maps display the relative amplitudes over 4 latency regions: 500-700, 700-900, 900-1100, and 1100-1400 msec.

Chapter 11

11 *General Discussion*

11.1 *Introductory remarks*

In six experiments the ERP correlates of recognition with and without retrieval of contextual information were investigated. One of the principal motivations for investigating the neural activity associated with recognition memory and context memory judgements was to assess the extent to which any putative ERP correlates of memory for a prior episode dissociated in a manner which was more consistent with the dual-process theory of recognition memory proposed by Jacoby and colleagues (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992), or the declarative memory hypothesis advocated by Squire (Squire, 1982a; Squire and Zola-Morgan, 1988), and Moscovitch (1992, 1994), among others.

The central section of this general discussion will focus on this issue, and will briefly review the relevant data, and expand upon the conclusions which have already been drawn. Prior to this section of the discussion, some aspects of the experimental data which have not been discussed to this point will be addressed.

11.2 *Misses and false alarms*

In five of the six experiments the ERPs to misses were compared to those to correct rejections. The analyses revealed no evidence for positive-going effects akin to those revealed by the comparisons of the ERPs to correctly recognised old and new words. Comparisons of the ERPs to false alarms and to correct rejections (experiments 1 to 4) also revealed no evidence for positive-going old/new effects. These findings are consistent with previous reports (Neville *et al.*, 1986; Rugg and Doyle, 1992; Smith, 1993), and suggest that neither the act of making an old response, nor the repetition of an old item, are sufficient in and of themselves to evoke old/new effects. The findings are therefore in accord with the view that the differences between ERPs evoked by correctly recognised old and new items are related to processes which support accurate report on the status of test stimuli.

In addition, it has been proposed that false alarms are made on the basis of fluency, where the level of fluency exceeds some criterion level above which stimuli are judged old (Jacoby, 1991; Yonelinas, 1994). To the extent that this interpretation is correct, the absence of any positive-going effects akin to those evident in the old/new effects for correctly recognised old items is consistent with the view that ERPs are insensitive to fluency. However, this is a relatively weak conclusion, since it is possible that the recorded ERPs were simply not sufficiently sensitive to detect the low levels of fluency associated with new items which were incorrectly judged old.

11.3 *Response confidence*

A second factor which has been proposed to reflect the differences between old and new items on tests of recognition memory is response confidence (Karis *et al.*, 1984). These proposals were applied only to the parietal old/new effect since they were based on studies in which ERPs were recorded for no more than a second, thereby making the detection of any frontal old/new effects unlikely. As has been noted (chapter 3), confidently detected stimuli are associated with more positive ERPs over the same time period in which old/new effects occur (Hillyard *et al.*, 1971; Ruchkin and Sutton, 1978). However, two previous studies of recognition memory report that the size of the old/new effect varies little as a function of confidence (Rugg *et al.*, 1995; Rugg and Doyle, 1992).

According to the dual-process view espoused by Jacoby and colleagues, recollection is an all-or-none process, one consequence of which is that recognition judgements made on the basis of recollection are always highly confident (Jacoby *et al.*, 1993; Yonelinas, 1994; see also Gardiner and Java, 1993). If this is not true of the other putative processes which support recognition judgements, then by this view there is a necessary confound between the levels of confidence associated with the responses comprising the hit/hit and hit/miss response categories.

Given that confidence ratings were not elicited in the experiments comprising this thesis, there is no direct evidence that the differences between the old/new effects for the hit/hit and hit/miss response categories are not in part due to variations in the confidence with which the initial old/new test judgements were made. However, the differences between the hit/hit and hit/miss old/new effects were at least as marked in experiment 4 as in experiment 3. In experiment 4 subjects had the option to respond don't know, which would presumably have attracted some unconfident old responses which might have contributed disproportionately to the hit/miss response category when the don't know response option was unavailable. Further, in experiments 1, 3 and 4 the RTs to correct old judgements did not predict the accuracy of the subsequent context judgement. To the extent that RT is related to response confidence, these findings suggest that the initial old judgements to these response categories were made equally confidently.

Of course, even if it were the case that there was reliable evidence for differences in response confidence across the hit/hit and hit/miss response categories, the differences between the hit/hit and hit/miss ERPs may not reflect this disparity. It is equally plausible to argue that the processes which support recognition with and without retrieval of context are associated with different levels of confidence, and it is the processes themselves which cause the differences between the ERPs associated with correct and incorrect context judgements.

These comments regarding response confidence are an example of a fundamental constraint that applies when making functional claims on the basis of ERP data. As was noted in chapter 2, given the present state of knowledge it is always possible to propose that differences between ERPs reflect processes that are merely correlated with, or are consequential upon, those processes that are the focus of interest in any given experiment.

11.4 *The negative shift in the ERPs to old judgements*

Prior to a general discussion and review of the antecedents and the possible functional significance of the parietal and frontal old/new effects, a third prominent modulation of the ERP waveforms revealed in this thesis will be discussed. This modulation consists of a negative-going shift in the ERPs to words judged old relative to words correctly judged new. The modulation onsets between 600 and 900 msec post-stimulus and is most evident at Pz. Examples of this negative shift can be seen in figure 7.1 (chapter 7), figure 9.1 (chapter 9), and figure 10.1 (chapter 10). A similar modulation is evident in some existing studies of recognition memory (Rugg *et al.*, 1995; Smith, 1993). As these previous studies recorded over an epoch of slightly less than one second in duration they revealed only the onset of this negative-going modulation. The studies in this thesis indicate that this negative shift continues until at least 1400 msec post-stimulus. However, the fact that these negative-going effects appear to be present on tasks in which subjects make only a single old/new judgement suggest that the effect is not solely related to the requirement to make a context judgement to words that have been judged old.

The significance of this negative-going shift is unclear. However, there is little evidence that it differentiates recognition which is accompanied or unaccompanied by retrieval of contextual information. In addition, negative-going shifts of similar magnitude were revealed by the comparisons of the ERPs to correct rejections and the ERPs to false alarms. This can be seen in figures 7.3, 8.3, and 9.4 (chapters 7, 8 and 9 respectively). These observations are consistent with the view that this negative shift is related more to the act of making an 'old' response than to processes which differentiate remembered and forgotten words.

Note that this view could not be advanced had ERPs to false alarms not been analysed. The practice of analysing ERPs to false alarms is not universally employed, often due to the fact that insufficient incorrect old judgements to new test items are made to permit formation of reliable averaged waveforms (e.g. Neville *et al.*, 1986; Smith, 1993). Even when false alarm ERPs are analysed, they are often associated with a low signal-to-noise ratio. This is of course in part a consequence of the fact that false alarm rates must be relatively low in order for performance on recognition memory tasks to be reliably above chance. None the less, there appears to be scope for a more systematic ERP investigation of the relationship between the processing which differentiates false alarms from genuinely remembered old words. This in turn may shed light on the significance of the negative-going shift to words attracting old judgements on tests of recognition memory.

11.5 *The processes supporting recognition with and without retrieval of context*

The discussion will now focus on the question of the processes which contribute to recognition memory, and memory for context. In experiments 3 and 4 two temporally overlapping and topographically distinct components were identified, both of which were modulated by recognition with and without retrieval of contextual information. The first of these - the parietal old/new effect - has previously been linked with processes supporting recollection (Paller and Kutas, 1992; Paller *et al.*, 1995; Smith, 1993; Smith and Halgren, 1989; Van Petten *et al.*, 1991), and with processes supporting relative fluency (Friedman, 1990; Johnson *et al.*, 1985; Rugg *et al.*, 1992; Rugg and Doyle, 1994).

11.51 *The parietal old/new effect and relative fluency*

The experiments reported in this thesis provide little support for the view that the parietal old/new effect is sensitive to relative fluency. The strongest evidence against a fluency interpretation stems from experiments 1 and 4, where the hit/hit old/new effect was reliably larger at parietal scalp sites than the hit/miss old/new effect. This pattern of results is inconsistent with the view that this effect is sensitive to fluency, irrespective of whether a relationship of exclusivity, redundancy, or independence obtains between the processes of fluency and recollection (see chapters 1 and 3).

As has been previously noted, further support for the view that the parietal old/new effect is insensitive to fluency stems from the fact that in experiments 1 and 2 there was no evidence that ERP old/new effects were larger when study and test modalities matched. To the extent that intra-modal priming is greater than inter-modal priming, an interpretation of the old/new effects in terms of relative fluency would have been supported if larger old/new effects were evident when study and test modality matched (Paller and Kutas, 1992; Wilding, Doyle and Rugg, 1995).

Finally, in experiment 6 the parietal old/new effect was larger for the spoken hit/hit ERPs than for the heard hit/hit ERPs. If the parietal old/new effect were sensitive to perceptual fluency then equivalent old/new effects would have been predicted in the two cases, given that at study and test all items were presented visually. The logic of this argument is similar to that employed by Paller and Kutas (1992, see chapter 3), who equated priming and observed a reliable parietal old/new effect only for words which had been studied in a semantic encoding task.

These findings constitute strong evidence that the parietal old/new effect does not index processes related to relative fluency, and therefore offer no direct support for the view that this process contributes to recognition memory performance. In experiment 3 there was some evidence for an early differentiation of the hit/hit and hit/miss ERPs at posterior scalp sites, where the hit/miss ERPs were more positive. As previously noted (chapter 7), the posterior distribution of these early effects is consistent with neuropsychological and neuroimaging data which link priming to regions of right occipital cortex (Fleischman *et al.*, 1995; Gabrieli *et al.*, 1995; Squire *et al.*, 1992). However, given the lack of evidence for equivalent early differentiations between the hit/hit and hit/miss ERPs in the other experiments in which ERPs to incorrect context judgements were analysed, the reliability of these effects in experiment 3 is not established (although see the caveats discussed in chapter 8).

Whilst the weight of evidence presented here suggests that ERPs are insensitive to relative fluency, the experimental findings of course cannot rule out the possibility that fluency does contribute to task performance, but the neural activity supporting judgements made on this basis cannot be reliably detected at the scalp. The findings of experiment 1 offer tentative support for this view. In this experiment the hit/miss ERPs for words presented auditorily at study were not reliably different from the ERPs to correct rejections. This was the only experiment in which a comparison of hit/miss and correct rejection ERPs revealed no evidence for a hit/miss old/new effect. The behavioural data in this experiment also indicated a large response bias to respond 'visual' for the context judgement, a finding which is consistent with the proposal that, in the absence of recollection, modality judgements can be made on the basis of relative fluency (Kelley *et al.*, 1989). By this view, fluent processing of a test item can be employed as the basis for a judgement that the item was previously encountered in the same modality.

If this interpretation is correct, then given that in experiment 1 all test items were presented visually, a proportion of responses made on this basis would comprise the auditory hit/miss response category. The responses comprising this category consist of an incorrect visual context judgement to correctly recognised old words. If fluency-based recognition is not indexed by ERPs, and was the basis for a proportion of the visual modality judgements made to words presented auditorily at study, then the likely effect of trials of this type would be to reduce the magnitude of any differences between the ERPs to correct rejections and those to incorrect modality judgements. This interpretation therefore supports the view that fluency contributes to recognition memory judgements, and offers to explain why the hit/miss ERPs in experiment 1 were not reliably different from those to correct rejections.

Converging evidence for this interpretation would stem from a ERP study of recognition memory in which the clarity of presentation of test items was varied. Behavioural studies using this approach report increased hit rates and false alarm rates for test items presented more clearly, which has been proposed to reflect the influence of fluent processing on recognition memory judgements (Jacoby and Whitehouse, 1989; Whittlesea *et al.*, 1990). If this interpretation of the behavioural findings is correct, and if ERPs are insensitive to fluency, then the size of any old/new effects observed should vary according to the clarity of presentation at test, with larger old/new effects to those words which were correctly judged old and presented less clearly.

11.52 *The parietal old/new effect and retrieval from declarative memory*

The experimental findings are consistent with the view that the parietal old/new effect is sensitive to the amount of information retrieved from memory (Rugg *et al.*, 1995). The strongest evidence for this view stems from experiment 4, where the introduction of the don't know option was intended clearly to distinguish between recollected and unrecollected words. In this experiment the hit/hit ERPs were reliably more positive than the hit/miss ERPs at parietal locations. The modulations of the parietal old/new effect in experiments 1 and 6 provide converging evidence for the view that the parietal effect varies with the quality or amount of retrieved information.

In experiment 3 there was no reliable evidence for a difference in the size of the parietal old/new effect for the hit/hit and the hit/miss ERPs. However, as has been previously discussed this likely reflects an attenuation of the parietal hit/hit old/new effect due to a proportion of trials comprising this response category which were not associated with retrieval of veridical contextual information.

Finally, in experiment 2 there was no evidence for a parietal asymmetry for either the hit/hit or the hit/miss ERPs. Over the latency region (400-800 msec) where the hit/miss ERPs were more positive than those to correct rejections the effects were not reliably different from those for the hit/hit ERPs. As noted in chapter 7, the possible influence of latency jitter on these auditory ERPs may have resulted in a decreased likelihood of observing differences in the morphology or the magnitude of any old/new effects. The time course of the hit/hit and hit/miss ERP old/new effects in experiment 2 suggests that these effects index at least some of the same processes which are reflected in the old/new effects observed in the other experiments comprising this thesis. However, given the paucity of studies investigating auditory recognition memory (for an exception see Domalski *et al.*, 1991), it is not clearly established whether part of the processing reflected by the parietal and frontal ERP old/new effects is specific to the modality of test presentation.

11.53 *The frontal old/new effect and memory for study context*

In comparison to the parietal old/new effect, the frontal effect has a more extended time course, and under certain experimental conditions an opposite asymmetry, being larger over the right hemisphere than over the left. In experiment 1 the recording epoch was just under 1 second, therefore precluding an analysis of any frontal old/new effects. In experiment 2 there was no evidence for a distinction between parietal and frontal old/new effects. As discussed above, the reasons for this are unresolved. Consequently, the following discussions of the frontal old/new effect are restricted to the results of experiments 3 to 6.

In experiments 3 and 4 the frontal effect was reliably larger for the hit/hit than for the hit/miss ERPs. Considered along with the findings that in experiment 4 the hit/miss ERPs showed a reliable frontal old/new effect, these results are consistent with the view that this effect is also sensitive to the quality or amount of information retrieved from memory. However, the results of experiments 5 and 6 suggest that the amount or quality of information retrieved from memory is not the sole determinant of the magnitude of the frontal old/new effect.

In experiment 5 the ERPs to the target hit response category were associated with a markedly asymmetric (right-greater-than-left) frontal old/new effect, whilst the ERPs to non-target hits showed no evidence for a frontal effect. In the discussion of these effects in chapter 9 it was concluded that the findings were consistent with the view the recollected words were processed differently depending upon whether they had been designated pre-experimentally as targets or non-targets, therefore linking the frontal old/new effect to task related factors.

This view was extended on the basis of the findings in experiment 6, where spoken and heard hit/hit ERPs were associated with equivalent frontal old/new effects, but the parietal old/new effect was reliably larger for words which had been spoken at study. For the interpretation of parietal old/new effects discussed above it was argued that there were good reasons to assume that words spoken at study are associated with more information which can be retrieved at test. The findings of experiment 6 are therefore inconsistent with the view that the frontal old/new effect is sensitive to solely the quality of information retrieved from memory. In the light of these findings it was proposed that the equivalent frontal effects reflect the fact that only a proportion of retrieved information may have been employed for the subsequent context judgement. As for the results of experiment 5, this interpretation links modulations of the frontal old/new effect to task-related factors.

11.54 *The relationship between parietal and frontal old/new effects*

The previous discussions of the processes indexed by the parietal and frontal old/new effects map fairly directly onto the processes which have been proposed to contribute to recognition with and without retrieval of context by Squire (Squire and Zola-Morgan, 1988), and Moscovitch (1992, 1994), amongst others. As has been noted, these models propose that two processes - a retrieval function and an integrative function - contribute to memory for context, and only the former is necessary for recognition memory judgements.

In one formulation of this model, Moscovitch (1992) described the processes contributing to memory for items and their context in terms of Fodor's distinction between modular and central systems (Fodor, 1983). He proposed that the retrieval function shared the properties of what Fodor termed input modules. The output of such modules is determined only by the input to them. That is, their function is reflexive. By this view, the retrieval function simply retrieves previously encoded information, irrespective of the task demands. The link between the parietal old/new effect and this retrieval function is supported by the findings that the size of the effect predicted the accuracy of context judgements in experiments 1 and 4. In experiment 2, memory for context was predicted by the duration of the parietal old/new effect. Finally, in experiment 6 the larger parietal old/new effect for spoken words has been interpreted above in terms of the amount of information retrieved from memory.

The second component of memory for context - the integrative function - is regarded as a component of a central system which is under conscious control. This function acts on the products of retrieval, and a direct implication of this definition is that any putative ERP correlate of such a process should be sensitive to factors other than the quality or amount of information retrieved, such as the particular task demands. The results of experiments 5 and 6 therefore support the link between the frontal old/new effect and the properties of this integrative function.

The model proposed by Moscovitch (1992) is intended as a general model of memory encoding and retrieval. It is therefore of considerable interest whether old/new effects similar to those observed on tests of recognition memory are observed on other direct tests of memory for a prior occurrence. In a recent study, Allan, Doyle, and Rugg (submitted) recorded ERPs in a modified word stem-cued recall task. In this task subjects were instructed to complete stems with previously studied words, or failing that to complete the stem with the first word that came to mind. In addition, following the completion subjects indicated - via an old/new judgement - whether they believed the completion was a word that had been presented at study.

ERPs evoked by stems which were completed with a studied word and which were correctly judged old showed an enhanced frontally maximal positivity in comparison to the ERPs evoked by stems completed with unstudied words. The distribution of this effect therefore provides some support for the view that it represents processing similar to that indexed by the frontal old/new effect in the modified tests of recognition memory reported here. It would be of considerable interest to investigate the ERP correlates of memory for study context on word stem cued recall tasks in order directly to assess the relationship between the effects observed on recall and recognition tasks.

11.6 *The neural generators of the parietal and frontal old/new effects*

It was noted in chapter 2 that considerable caution is necessary when making inferences about the neural generators of scalp-recorded ERP waveforms. However, the previous discussions have linked the parietal and frontal old/new effects to the processes of retrieval from declarative memory, and the integration of the retrieved information into a coherent representation of a prior learning episode. These processes have in turn been linked with the integrity of the medial temporal lobes and frontal lobes respectively (Moscovitch, 1992; Moscovitch, 1994; Squire and Knowlton, 1994; a more detailed account of this framework and the relevant neuropsychological findings was provided in chapter 1). In light of these proposals it is of interest whether the available data suggest a link between these brain structures and the scalp-recorded ERPs reported here. The parietal and frontal old/new effects will be discussed in turn.

11.61 *Generators of the parietal old/new effect*

The previous discussions have linked this effect to the process of retrieval of information from the declarative memory system, implying that the effect is in some way reflective of neural activity in the medial temporal lobes. There are two principal lines of evidence which bear on this issue. The first stems from studies investigating the neural basis of the P300, which is maximal over centro-parietal scalp regions, and is typically observed in the same time window over which the parietal old/new effect is evident (for a brief review of the P300 component see chapter 2). Intracranial recordings in the medial-temporal lobes and associated structures during oddball tasks (see chapter 2) have revealed enhanced potentials to target stimuli in comparison to non-targets (Halgren, Squires, Wilson, Rohrbaugh, Babb and Crandall, 1980; McCarthy, Wood, Williamson and Spencer, 1989). However, scalp recorded ERPs in patients with lesions of the medial temporal lobes exhibit P300 effects which do not differ markedly from those in intact subjects (Johnson, 1989; Rugg, 1995; Stapleton, Halgren and Moreno, 1987; Wood, McCarthy, Allison, Goff, Williamson and Spencer, 1982). These findings suggest that ERP activity generated locally within the hippocampus and adjacent structures is not volume conducted to the scalp.

Whilst these findings suggest that the parietal old/new effect is not a direct reflection of activity in medial-temporal lobe structures, it is nonetheless of interest that intra-cerebral ERP recordings from medial temporal lobe structures during recognition memory tasks differentiated between items correctly judged old and items correctly judged new in the same time frame over which parietal old/new effects are evident at the scalp (Heit, Smith and Halgren, 1990; Smith, Stapleton and Halgren, 1986). Together with findings that unilateral lesions involving the medial temporal lobes abolish the parietal old/new effect (Rugg, Roberts, Potter, Pickles and Nagy, 1991; Smith and Halgren, 1989), the intra-cerebral recordings support the view that, if not generated directly in the medial temporal lobes, the old/new effect is dependent upon the functional integrity of this region, and may serve as an electrophysiological index of medial temporal lobe function.

11.62 *Generators of the frontal old/new effect*

In experiments 3 to 6, the differences between the hit/hit and correct rejection ERPs post-600 msec were most sustained at frontal electrode sites, and were larger over the right hemisphere than over the left. This pattern of activity is at least consistent with the view that the neural generators of this effect are located in the frontal lobes. Two lines of evidence support this position. First, given that the frontal old/new effects have been linked to processes contributing to memory for context, the idea that the effect has a locus in the frontal lobes is consistent with findings that patients with frontal lesions have disproportionately poor memory for context or source (Janowsky *et al.*, 1989; Schacter *et al.*, 1984; Shimamura and Squire, 1987), and that disproportionately poor memory for source is correlated with poor performance on behavioural tests thought to assess frontal function (Glisky *et al.*, 1995; Parkin and Walter, 1992).

Second, a series of PET studies have revealed activation in right prefrontal cortex on tasks requiring retrieval from episodic memory (Buckner, Petersen, Ojemann, Miezin and Squire, 1995; Buckner and Tulving, 1995; Fletcher, Frith, Grasby, Shallice, Frackowiack and Dolan, 1995; Haxby, Horwik, Maisog, Ungerleider, Mishkin, Schapiro, Rapoport and Grady, 1993; Kapur, Craik, Tulving, Wilson, Houle and Brown, 1994; Tulving, Kapur, Craik, Moscovitch and Houle, 1994a; Tulving, Kapur, Markowitsch, Craik, Habib and Houle, 1994b). With three exceptions these studies have not directly distinguished between retrieval attempts and retrieval success. Tulving *et al.*. (1994b) compared the PET correlates of recognition for spoken sentences in one test condition where all test words were old, and a second in which all test words were new. In this study, right prefrontal activation was revealed only in the 'all old' condition, a finding which suggests a link between prefrontal activation and retrieval success. However, the results of this study are open to an alternative interpretation, since subjects were aware in advance of whether the subsequent test phase would consist of a high or low proportion of old words. Consequently, the activations observed may have been modulated by different strategies that subjects adopted in the two experimental conditions.

The findings in this study also contrast with those of Kapur *et al.*. (1994), where subjects took part in two recognition memory tasks in which the proportions of old and new items were 15:85 and 85:15 respectively. Both recognition tasks resulted in activation of right prefrontal cortex, which was not distinguishable between the two tasks. On the basis of these findings the authors concluded that the right prefrontal activation was more concerned with retrieval attempts than with the processing of successfully retrieved information.

Rugg and colleagues (Rugg, Fletcher, Frith, Frackowiack and Dolan, submitted) recently addressed the issue of the functional significance of the right prefrontal activation in PET studies of memory retrieval, and replicated the findings of Kapur *et al.*. (1994) for the comparisons of the PET correlates of recognition memory where there was either a high or a low proportion of old items in the test conditions. However, Rugg *et al.* also included a test condition in which there were no old items. Subtraction of the PET images for this condition from the high and low density conditions revealed right prefrontal activation only in those conditions in which there was a proportion of old items. These findings therefore suggest that the frontal activation revealed in PET studies of recognition memory is not solely related to retrieval attempts.

These findings converge with those of the present study to suggest that neural activity in prefrontal cortex supports retrieval of contextual information. The view that the right frontal activity is not due solely to retrieval attempts is consistent with the findings that the ERPs to false alarms show no evidence for frontal old/new effects, and the old/new effects for the hit/miss ERPs are smaller than those for the hit/hit ERPs.

In addition to these conclusions, the differences in the magnitude of the hit/hit and hit/miss frontal old/new effects in experiments 3 and 4 support the view that activity in right prefrontal cortex varies with the amount of information retrieved. Further, the modulations of the frontal old/new effect in experiments 5 and 6 suggest that neural activity in this region should also be modulated by task demands. No PET study to date has directly investigated memory for study context, and it remains to be seen whether data obtained from other imaging modalities supports these conclusions regarding the activity in the frontal lobes on tests of memory which explicitly require retrieval of study context.

The rapid development of brain imaging techniques such as PET and fMRI demonstrates enormous potential for mapping the relationship between neural structures and cognition. At the same time, the limitations of these techniques - in particular the fact that they provide an aggregate measure of neural activity over a period of several seconds - have served to emphasise the utility of event-related potential research, given the millisecond resolution of this electrophysiological index of brain function. The principle of employing data from more than one imaging modality on equivalent tasks as sources of converging evidence is sound, offering as it does to speak both to the issue of the neural structures instantiating cognition, as well as the temporal relationship between the neural activity in different brain structures. How fruitful this approach will be in practice is still very much an open question.

11.7 *Additional considerations*

In the final sections below some additional considerations concerning the generality of the experimental results reported here are introduced, and some corresponding implications for theories of recognition with and without retrieval of context are entertained.

11.71 *The use of words as stimuli*

In all six experiments stimuli consisted of low- frequency words, and in general there has been relatively little investigation of the ERP correlates of memory for non-verbal stimuli on test of recognition memory (for two exceptions see Friedman, 1990; Noldy, Stelmack and Campbell, 1990). To date there are no published studies of retrieval of contextual information associated with non-verbal stimuli. Consequently, the extent to which the ERP old/new effects observed here are specific to words is unresolved. The issue is obviously of some importance, since a comparison of old/new effects evoked by verbal and non-verbal stimuli would permit further delineation of the processes which the parietal and frontal old/new effects do in fact index. For example, if equivalent parietal old/new effects were observed for correctly recognised pictures and correctly recognised words this finding would constrain functional interpretations of the type and/or stage of processing which the parietal old/new effect represents.

11.72 *Retrieval of different forms of contextual information*

The final experiment reported in this thesis investigated the sensitivity of the frontal and parietal old/new effects to the retrieval of different kinds of contextual information. The findings built on those of experiments 3, 4, and 5, in which the context discrimination was between words spoken in either a male or a female voice. Whilst the findings of experiment 6 extended our knowledge regarding the factors influencing the parietal and frontal old/new effects, it is arguable whether these effects still principally reflected contextual information about speaker voice, since words were either spoken or heard at study.

To assess the generality of these effects further studies are required where different contextual manipulations are employed (for example, see Baddeley, 1982). In a recent example of this type, Schloerscheidt *et al.*. (Schloerscheidt, Rugg, Doyle, Cox and Patching, 1995) recorded ERPs in a visual recognition memory test following a study phase in which a pair of words were presented, and subjects were required to generate a sentence containing the words. In the subsequent test phase subjects made old/new judgements to visually presented words, and were asked to recall the associated studied word. The hit/hit ERPs (words correctly judged old for which the associate was correctly retrieved) were associated with a temporally extended parietal old/new effect, but no reliable frontal old/new effect.

These findings are consistent with view that the duration of the parietal old/new effect varies with the type of information retrieved from memory. In addition, the marked absence of a frontal old/new effect suggests that the processing indexed by this effect is specific to certain kinds of contextual information. Considered jointly, these preliminary findings indicate that there is considerable scope for further investigations of the stimulus and task parameters which modulate the parietal and frontal old/new effects on tests of retrieval of contextual information from long-term memory.

11.73 *Graded recollection*

The concept of graded recollection was discussed briefly in the introduction to the relevant behavioural memory literature in chapter 1, and it has been proposed that both parietal and frontal old/new effects are sensitive to the amount or the quality of information that is retrieved from memory. This view is counter to that proposed by Jacoby and colleagues (Jacoby *et al.*, 1993; Yonelinas, 1994); for whom recollection is an all-or-none process, as was noted when discussing the relationship between recollection and response confidence.

It appears that there are two forms that graded recollection can take, and the distinction between them has not been clearly specified in the literature. In one sense graded, or *partial*, recollection can refer to the retrieval of only some aspects of a prior episode. This definition acknowledges that, on tasks such as those comprising this thesis, it is in principle possible to recollect the prior study episode, but be unable to make the required context discrimination. This form of graded recollection is articulated in the source monitoring framework due principally to Johnson and colleagues (Johnson *et al.*, 1993; Johnson and Raye, 1981). According to this framework, different forms of contextual information can be retrieved depending upon the particular learning episode, the operations performed at the time of encoding, and the operations performed at retrieval.

The second form that graded recollection can take is variation in the quality of information that is retrieved concerning a specific contextual attribute of a prior episode. These two conceptualisations of graded recollection are of course not mutually exclusive. However, the distinction between the two forms is important since they make different predictions regarding the bases for task judgements. By the former view, recollection of a prior episode only supports context judgements for those contextual details that were retrieved. By the latter view, even if relevant contextual information is retrieved, it is still possible that the quality of the information is not sufficient to make an accurate context judgement, depending upon the nature of the context discrimination that is required. Presumably, any attempts to model recollection using a form of signal-detection theory analysis must adopt a definition of recollection which incorporates the notion of a gradation in the quality of contextual information that is retrieved from long-term memory.

In summary, it seems reasonable to assume that the conception of recollection as an all-or-none process is an over-simplification. However, it is not clearly established how recollection should be described. The fact that both the parietal and frontal old/new effects reported in this thesis are modulated according to the accuracy of context judgements suggests that ERP studies of recognition with and without retrieval of study context can provide further information which will speak to these issues.

11.8 *Concluding Remarks*

Six experiments investigated the ERP correlates of memory for words and for the context in which they were presented. The principal findings were that two distinct processes contribute to recognition which is accompanied or unaccompanied by retrieval of study context. There was no reliable evidence that the processes supporting these two forms of explicit memory were qualitatively different. Contrary to the view espoused by some dual-process theorists (Jacoby and Dallas, 1981; Jacoby and Kelley, 1992), the findings are therefore consistent with the view that the relationship between recognition with and without retrieval of context is one of degree rather than one of kind.

These findings were interpreted within the extensions of the declarative memory hypothesis which have been most directly associated with the work of Squire (1994; Squire and Zola-Morgan, 1988), and Moscovitch (1992; 1994). Converging evidence from neuropsychological studies, and studies of memory using other imaging techniques, was employed in order to explore the relationship between the processing indexed by event-related potentials on tasks of memory retrieval, and the brain structures on whose integrity these processes depend. The results contribute to existing knowledge regarding the neural and functional processes which contribute to recognition with and without retrieval of contextual information.

References

Allan, K., Doyle, M. C., and Rugg, M. D. An event-related potential study of word-stem cued recall. *submitted.*

Allison, T., Wood, C. C., and McCarthy, G. The central nervous system. In *Psychophysiology: Systems, Processes, and Applications.* M. G. H. Coles, S. W. Porges, and E. Donchin (Eds), pp. 5-25. Guilford Press, New York, 1986.

Anderson, J. R., and Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review* **79**, 97-123, 1972.

Atkinson, R. C., and Juola, J. F. Factors influencing speed and accuracy of word recognition. In *Attention and Performance IV.* S. Kornblum (Ed), pp. 583-612. Academic Press, New York, 1973.

Atkinson, R. C., and Juola, J. F. Search and decision processes in recognition memory. In *Contemporary Developments in Mathematical Psychology.* D. H. Krantz, R. C. Atkinson, and P. Suppes (Eds), pp. 243-293. W.H. Freeman, San Francisco, 1974.

Atkinson, R. C., and Shiffrin, R. M. Human memory: A proposed system and its control processes. In *The psychology of learning and motivation.* K. Spence, and J. Spence (Eds), pp. 89-195. Academic Press, New York, 1968.

Baddeley, A. D. Domains of recollection. *Psychological Review* **89**, 708-729, 1982.

Baddeley, A. D., and Hitch, G. Working memory. In *Recent Advances in Learning and Motivation.* G. A. Bower (Ed), pp. 47-90. Academic Press, New York, 1974.

Barrett, S. E., and Rugg, M. D. Event-related potentials and the phonological matching of pictures. *Psychophysiology* **30**, 327-339, 1990.

Batchelder, W. H., and Riefer, D. M. Multinomial processing models of source monitoring. *Psychological Review* **97**, 548-564, 1990.

Bentin, S., McCarthy, G., and Wood, C. C. Event-related potentials, lexical decision, and semantic priming. *Electroencephalography and Clinical Neurophysiology* **60**, 343-355, 1985.

Bentin, S., Moscovitch, M., and Heth, I. Memory with and without awareness: performance and electrophysiological evidence of savings. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18**, 1270-1283, 1992.

Bentin, S., and Peled, B. S. The contribution of task-related factors to ERP repetition effects at short and long lags. *Memory and Cognition* **18(4)**, 359-366, 1990.

Berg, P., and Scherg, M. A fast method for computation of multiple-shell head models. *Electroencephalography and Clinical Neurophysiology* **90**, 58-64, 1994.

Binnie, C. D. Recording techniques: Montages, electrodes, amplifiers, and filters. In *A Textbook of Clinical Neurophysiology.* A. M. Halliday, S. R. Butler, and R. Paul (Eds), pp. 3-22. Wiley, New York, 1987.

Bowers, J. S., and Schacter, D. L. Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **16(3)**, 404-416, 1990.

Brown, J. An analysis of recognition and recall and of problems in their comparison. In *Recall and Recognition.* J. Brown (Ed), pp. 1-36. Wiley, London, 1975.

Buchner, A., Erdfelder, E., Vaterrodt-Plünnecke, B. Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General* **124(2)**, 137-160, 1995.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., and Squire, L. R. Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience* **15**, 12-29, 1995.

Buckner, R. L., and Tulving, E. Neuroimaging studies of memory: Theory and recent PET results. In *Handbook of Neuropsychology, Volume 10.* F. Boller, and J. Grafman (Eds), pp. 439-466. Elsevier, Amsterdam, 1995.

Clarke, R., and Morton, J. Cross-modality facilitation in tachistoscopic word recognition. *Quarterly Journal of Experimental Psychology* **35A**, 79-96, 1983.

Cofer, C. N. Conditions for the use of verbal associations. *Psychological Bulletin* **68**, 1-12, 1967.

Cohen, N. J., and Squire, L. R. Preserved learning and retention of a pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science* **210**, 207-210, 1980.

Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, T. W., and Donchin, E. A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance* **11**, 529-553, 1985.

Coles, M. G. H., Gratton, G., and Fabiani, M. Event-related brain potentials. In *Principles of psychophysiology: physical, social and inferential elements.* J. T. Cacciopo, and L. G. Tassinary (Eds), pp. 413-455. Cambridge University Press, Cambridge, 1990.

Cooper, R., Osselton, J. W., and Shaw, J. C. *EEG Technology*, Butterworths, London, 1980.

Curran, T., and Hintzman, D.L. Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **21(3)**, 531-547, 1995.

Domalski, P., Smith, M. E., and Halgren, E. Cross-modal repetition effects on the N4. *Psychological Science* **2(3)**, 173-178, 1991.

Donchin, E., and Coles, M. G. H. Is the P300 component a manifestation of context updating ? *Behavioral and Brain Sciences* **11**, 355-372, 1988.

Donchin, E., Karis, D., Bashore, T. R., Coles, M. G. H., and Gratton, G. Cognitive psychophysiology and human information processing. In *Psychophysiology: Systems, Processes, and Applications.* M. G. H. Coles, E. Donchin, and S. E. Porges (Eds), pp. 702-718. Guilford Press, New York, 1986.

Donchin, E., Ritter, W., and McCallum, W. C. Cognitive psychophysiology: The endogenous components of the ERP. In *Event-Related Potentials in Man.* E. Callaway, P. Tueting, and S. H. Koslow (Eds), pp. 349-442. Academic Press, London, 1978.

Dorfman, J., Kihlstrom, J. F., Cork, R. C., and Misiaszek, J. Priming and recognition in ECT-induced amnesia. *Psychonomic Bulletin and Review* **2**, 244-248, 1995.

Dosher, B. A. Discriminating pre-experimental (semantic) from learned (episodic) associations. *Cognitive Psychology* **16**, 519-555, 1984.

Duncan-Johnson, C. C., and Donchin, E. On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology* **14**, 456-467, 1977.

Dunn, J. C., and Kirsner, K. Implicit memory: Task or process ? In *Implicit Memory: Theoretical Issues.* S. Lewandowsky, J. C. Dunn, and K. Kirsner (Eds), pp. 17-31. Erlbaum, Hillsdale, New Jersey, 1989.

Fabiani, M., Karis, D., and Donchin, E. P300 and recall in an incidental memory paradigm. *Psychophysiology* **23**, 298-308, 1986.

Fleischman, D. A., Gabrieli, J. D. E., Reminger, S., Rinaldi, J., and Morrell, F. Conceptual priming in perceptual identification for patients with Alzheimers disease and a patient with right occipital lobectomy. *Neuropsychology* **33**, 18-25, 1995.

Fletcher, P. C., Frith, C. D., Grasby, P. M., Shallice, T., Frackowiack, R. S. J., and Dolan, R. J. Brain systems for encoding and retrieval of auditory-verbal memory. *Brain* **118**, 401-416, 1995.

Fodor, J. *The modularity of mind*, MIT/Bradford Press, Cambridge, Massachusetts, 1983.

Friedman, D. Cognitive event-related potential components during continuous recognition memory for pictures. *Psychophysiology* **27**, 136-148, 1990.

Friedman, D. ERPs during continuous recognition memory for words. *Biological Psychology* **30**, 61-88, 1990.

Friedman, D., and Sutton, S. Event-related potentials during continuous recognition memory. In *Current Trends in Event-Related Potential Research. (Supplement 40 to Electroencephalography and Clinical Neurophysiology).* R. Johnson, J. W. Rohrbaugh, and R. Parasuraman (Eds), pp. 316-321. Elsevier, Amsterdam, 1987.

Gabrieli, J. D. E., Fleischman, D. A., Keane, M. M., Reminger, S. L., and Morrell, F. Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science* **6**, 76-82, 1995.

Gaffan, D. Recognition impaired and association intact in the memory of monkeys after transection of the fornix. *Journal of Comparative and Physiological Psychology* **86**, 1100-1109, 1974.

Gardiner, J. M. Functional aspects of recollective experience. *Memory and Cognition* **16**, 309-313, 1988.

Gardiner, J. M., and Java, R. I. Recollective experience in word and non-word recognition. *Memory and Cognition* **18**, 23-30, 1990.

Gardiner, J. M., and Java, R. I. Recognising and Remembering. In *Theories of Memory.* A. Collins, M. A. Conway, S. E. Gathercole, and P. E. Morris (Eds), pp. 163-188. Erlbaum, Hillsdale, New Jersey, 1993.

Gardiner, J. M., and Parkin, A. J. Attention and recollective experience in recognition memory. *Memory and Cognition* **18**, 579-583, 1990.

Gardner, H., Boller, F., Moreines, J., and Butters, N. Retrieving information from Korsakoff patients: Effects of categorical cues and reference to the task. *Cortex* **9**, 165-175, 1973.

Gillund, G., and Shiffrin, R. M. A retrieval model for both recognition and recall. *Psychological Review* **91**, 1-67, 1984.

Glisky, E. L., Polster, M. L., and Routhieaux, B. C. Double dissociation between item and source memory. *Neuropsychology* **9**, 229-235, 1995.

Gorman, A. M. Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology* **61**, 23-29, 1961.

Graf, P., and Komatsu, S. Process dissociation procedure: Handle with caution ! *European Journal of Cognitive Psychology* **6**, 113-129, 1994.

Graf, P., Mandler, G., and Haden, P. E. Simulating amnesic symptoms in normals. *Science* **218**, 1243-1244, 1982.

Graf, P., and Schacter, D. L. Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11**, 501-518, 1985.

Graf, P., Squire, L. R., and Mandler, G. The information that amnesics do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **10**, 164-178, 1984.

Green, D. M., and Swets, J. A. *Signal detection theory and psychophysics*, Wiley, New York, 1966.

Gregg, V. H., and Gardiner, J. H. Recognition memory and awareness: A large effect of study-test modalities on 'Know' responses following a highly perceptual orienting task. *European Journal of Cognitive Psychology* **6**, 131-147, 1994.

Haist, F., Shimamura, A. P., and Squire, L. R. On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18**, 691-702, 1992.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., and Crandall, P. H. Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science* **210**, 803-805, 1980.

Haxby, J. V., Horwik, B., Maisog, J. M., Ungerleider, L. G., Mishkin, M., Schapiro, M. B., Rapoport, S. I., and Grady, C. L. Frontal and temporal participation in long-term recognition memory for faces: a PET-rCBF activation study. *Journal of Cerebral Blood Flow Metabolism* **13(Supplement 1)**, S499, 1993.

Hayman, C. A. G., and Tulving, E. Contingent dissociation between recognition and fragment completion: The method of triangulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 228-240, 1989.

Heit, G., Smith, M. E., and Halgren, E. Neuronal activity in the human medial temporal lobe during recognition memory. *Brain* **113**, 1093-1112, 1990.

Hillyard, S. A., and Kutas, M. Electrophysiology of cognitive processing. *Annual Review of Psychology* **34**, 33-61, 1983.

Hillyard, S. A., and Picton, T. W. Electrophysiology of cognition. In *Handbook of Physiology, Section 1: Neurophysiology.* F. Plum (Ed), pp. 519-584. American Physiological Society, New York, 1987.

Hillyard, S. A., Squires, K. C., Bauer, J. W., and Lindsay, P. H. Evoked potential correlates of auditory signal detection. *Science* **172**, 1357-1360, 1971.

Hintzman, D. L. Simpson's paradox and the analysis of memory retrieval. *Psychological Review* **87**, 398-410, 1980.

Hintzman, D. L., and Curran, T. Retrieval dynamics of recognition memory and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language* **33**, 1-18, 1994.

Hirsh, R. The hippocampus and contextual retrieval of information from memory: A theory. *Behavioural Psychology* **12**, 421-444, 1974.

Hirst, W., Johnson, M. K., Phelps, E. A., Risse, G., and Volpe, B. T. Recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 445-451, 1986.

Hirst, W., Johnson, M. K., Phelps, E. A., and Volpe, B. T. More on recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 758-762, 1988.

Humphreys, M. S., Bain, J. D., and Pike, R. Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review* **96**, 208-233, 1989.

Hyde, T. S., and Jenkins, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behaviour* **12**, 471-480, 1973.

Jacoby, L. L. Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **9**, 21-38, 1983.

Jacoby, L. L. Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behaviour* **22**, 485-508, 1983.

Jacoby, L. L. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language* **30**, 513-541, 1991.

Jacoby, L. L., Allan, L. G., Collins, J. C., and Larwill, L. K. Memory influences subjective experience: Noise judgements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 240-247, 1988.

Jacoby, L. L., and Dallas, M. On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General* **3**, 306-340, 1981.

Jacoby, L. L., and Kelley, C. Unconscious influences of memory: Dissociations and automaticity. In *The Neuropsychology of Consciousness.* A. D. Milner, and M. D. Rugg (Eds), pp. 201-233. Academic Press, London, 1992.

Jacoby, L. L., Kelley, C. M., and Dywan, J. Memory attributions. In *Varieties of Memory and Consciousness.* H. L. Roediger, and F. M. Craik (Eds), pp. 391-422. Erlbaum, Hillsdale, New Jersey, 1989.

Jacoby, L. L., Toth, J. P., and Yonelinas, A. P. Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General* **122**, 139-154, 1993.

Jacoby, L. L., Toth, J. P., Yonelinas, A. P., and Debner, J. A. The relationship between conscious and unconscious influences: Independence or redundancy ? *Journal of Experimental Psychology: General* **2**, 216-219, 1994.

Jacoby, L. L., and Whitehouse, K. An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General* **118**, 126-135, 1989.

Jacoby, L. L., and Witherspoon, D. Remembering without awareness. *Canadian Journal of Psychology* **32**, 300-324, 1982.

Jacoby, L. L., Woloshyn, V., and Kelley, C. L. Becoming famous without being recognised: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General* **118**, 115-125, 1989.

Jacoby, L. L., Yonelinas, A. P., and Jennings, J. M. The relationship between conscious and unconscious (automatic) influences: A declaration of independence. In *Scientific approaches to the question of consciousness: The 25th annual Carnegie symposium on cognition.* J. Cohen, and J. W. Schooler (Eds), Erlbaum, Hillsdale, New Jersey, *in press*.

Janowsky, J. S., Shimamura, A. P., and Squire, L. R. Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia* **27**, 1043-1056, 1989.

Jasper, H. A. The ten-twenty system of the international federation. *Electroencephalography and Clinical Neurophysiology* **10**, 371-375, 1958.

Jetter, W., Poser, U., Freeman, R. B., and Markowitsch, J. H. A verbal long-term memory deficit in frontal lobe damaged patients. *Cortex* **22**, 229-242, 1986.

John, E. R., Ruchkin, D. S., and Vidal, J. J. Measurement of event-related potentials. In *Event-Related Brain Potentials in Man.* E. Callaway, P. Tueting, and S. H. Koslow (Eds), pp. 93-151. Academic Press, New York, 1978.

Johnson, M. K., Foley, M. A., and Leach, K. The consequences for memory of imagining in another persons voice. *Memory and Cognition* **16**, 337-342, 1988.

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. Source monitoring. *Psychological Bulletin* **114**, 3-28, 1993.

Johnson, M. K., Kounios, J., and Reeder, J. A. Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20**, 1409-1419, 1994.

Johnson, M. K., and Raye, C. L. Reality monitoring. *Psychological Review* **88**, 67-85, 1981.

Johnson, M. K., Raye, C. L., Foley, H. J., and Foley, M. A. Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology* **94**, 37-64, 1981.

Johnson, R. Auditory and visual P300s in temporal lobectomy patients: Evidence for modality dependent generators. *Psychophysiology* **26**, 633-650, 1989.

Johnson, R. On the neural generators of the P300 component of the event-related potential. *Psychophysiology* **30**, 94-101, 1993.

Johnson, R. Event-related potential insights into the neurobiology of memory systems. In *Handbook of Neuropsychology, Volume 10.* J. C. Baron, and J. Grafman (Eds), pp. 135-164. Elsevier, Amsterdam, 1995.

Johnson, R., Pfefferbaum, A., and Kopell, B. S. P300 and long-term memory: Latency predicts recognition performance. *Psychophysiology* **22**, 497-507, 1985.

Johnston, W. A., Dark, V., and Jacoby, L. L. Perceptual fluency and recognition judgements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11**, 3-11, 1985.

Johnston, W. A., Hawley, K. J., and Elliot, J. M. G. Contribution of perceptual fluency to recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **17**, 210-223, 1991.

Jones, G. V. Independence and exclusivity among psychological processes: Implications for the structure of recall. *Psychological Review* **94**, 229-235, 1987.

Joordens, S., and Merikle, P. M. Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General* **122**, 462-467, 1993.

Kapur, S., Craik, F. I. M., Tulving, E., Wilson, M., Houle, S., and Brown, G. M. Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Science* **91**, 2012-2015, 1994.

Karis, D., Fabiani, M., and Donchin, E. P300 and memory: Individual differences in the Von Restorff effect. *Cognitive Psychology* **16**, 177-216, 1984.

Kelley, C. M., Jacoby, L. L., and Hollingshead, A. Direct vs. indirect tests of memory for source: Judgements of modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 1101-1108, 1989.

Keselman, H. J., and Rogan, J. C. Repeated measures F-tests and psychophysiological research: Controlling the number of false positives. *Psychophysiology* **17**, 499-503, 1980.

Kinsbourne, M., and Wood, F. Short term memory and pathological forgetting. In *Short term memory.* D. Deutsch, and A. J. Deutsch (Eds), pp. 249-284. Academic Press, New York, 1975.

Kintsch, W. Memory and decision aspects of recognition learning. *Psychological Review* **74**, 496-504, 1967.

Kintsch, W. *Learning, Memory, and Conceptual Processes,* Wiley, New York, 1970.

Kirsner, K., Milech, D., and Standen, P. Common and modality specific processes in the mental lexicon. *Memory and Cognition* **11**, 621-630, 1983.

Kirsner, K., and Smith, M. Modality effects in word identification. *Memory and Cognition* **2**, 637-640, 1974.

Knowlton, B. J., and Squire, L. R. Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **21**, 699-710, 1995.

Kopelman, M. D., and Corn, T. H. Cholinergic 'blockade' as a model for cholinergic depletion. A comparison of the memory deficits with those of Alzheimer-type dementia and the Alcoholic Korsakoff syndrome. *Brain* **111**, 1079-1110, 1988.

Kucera, H., and Francis, W. N. *Computational analysis of present-day American English,* Brown University Press, Providence, RI, 1967.

Kutas, M. Review of event-related potential studies of memory. In *Perspectives in Memory Research.* M. S. Gazzaniga (Ed), pp. 182-217. MIT Press, Cambridge, Massachusetts, 1988.

Kutas, M., and Hillyard, S. A. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* **207**, 203-205, 1980.

Kutas, M., and Hillyard, S. A. The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia* **20**, 579-590, 1982.

Kutas, M., and Hillyard, S. A. Brain potentials during reading reflect word expectancy and semantic association. *Nature* **307**, 161-163, 1984.

Kutas, M., McCarthy, G., and Donchin, E. Augmenting mental chronometry: The P300 as a measure of evaluation time. *Brain Research* **202**, 95-115, 1987.

Light, L. L., and Singh, A. Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 531-541, 1987.

Lins, O. G., Picton, T. W., Berg, P., and Scherg, M. Ocular artifacts in EEG and event-related potentials. 1. Scalp Topography. *Brain Topography* **6**, 51-63, 1993.

Mandler, G. Recognising: The judgment of previous occurrence. *Psychological Review* **87**, 252-271, 1980.

Mandler, G., Goodman, G. O., and Wilkes-Gibbs, D. L. The word-frequency paradox in recognition. *Memory and Cognition* **10**, 33-42, 1982.

Marslen-Wilson, W. Function and process in spoken word recognition. In *Attention and Performance X.* H. Bouma, and D. Bouhuis (Eds), pp. 125-150. Erlbaum, Hillsdale, New Jersey, 1984.

Mayes, A. R. *Human organic memory disorders*, Cambridge University Press, Cambridge, 1988.

Mayes, A. R. Automatic memory processes in amnesia: How are they mediated ? In *The Neuropsychology of Consciousness.* A. D. Milner, and M. D. Rugg (Eds), pp. 235-261. Academic Press, London, 1992.

McCarthy, G., and Wood, C. C. Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology* **62**, 203-208, 1985.

McCarthy, G., Wood, C. C., Williamson, P. D., and Spencer, D. D. Task-dependent field potentials in the human hippocampal formation. *Journal of Neuroscience* **9**, 453-468, 1989.

McKoon, G., Ratcliff, R., and Dell, G. S. A critical evaluation of the semantic-episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 295-306, 1986.

Milner, B. Amnesia following operation on the temporal lobes. In *Amnesia.* C. W. M. Whitty, and O. M. Zangwill (Eds), pp. 109-133. Butterworths, London, 1966.

Milner, B. Interhemispheric differences in the location of psychological processes in man. *British Medical Bulletin* **27**, 272-277, 1971.

Milner, P. A cell assembly theory of hippocampal amnesia. *Neuropsychologia* **27**, 31-40, 1989.

Morris, C. D., Bransford, J. D., and Franks, J. J. Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behaviour* **16**, 519-533, 1977.

Moscovitch, M. Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience* **4**, 257-267, 1992.

Moscovitch, M. Models of consciousness and memory. In *The Cognitive Neurosciences.* M. S. Gazzaniga (Ed), pp. 1341-1356. MIT Press, Cambridge, Massachusetts, 1994.

Mulligan, N., and Hirshman, E. Speed accuracy trade-offs and the dual-process model of recognition memory. *Journal of Memory and Language* **34**, 1-18, 1995.

Näätänen, R. Processing negativity: An evoked-potential reflection of selective attention. *Psychological Bulletin* **92**, 605-640, 1982.

Näätänen, R., and Picton, T. W. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology* **24**, 375-425, 1987.

Neisser, U. An experimental distinction between perceptual processes and verbal response. *Journal of Experimental Psychology* **47**, 399-402, 1954.

Neville, H. J., Kutas, M., Chesney, G., and Schmidt, A. L. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language* **25**, 75-92, 1986.

Noldy, N. E., Stelmack, R. M., and Campbell, K. B. Event-related potentials and recognition memory for pictures and words: The effects of intentional and incidental learning. *Psychophysiology* **27**, 417-428, 1990.

Norman, D. A., and Wickelgren, W. A. Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology* **6**, 192-208, 1969.

Nunez, P. L. *Electric fields of the brain: The neurophysics of EEG*, Oxford University Press, New York, 1981.

Nunez, P. L. Physical principles and neurophysiological mechanisms underlying event-related potentials. In *Event-Related Brain Potentials: Basic Issues and Applications.* J. W. Rohrbaugh, R. Parasuraman, and R. Johnson (Eds), pp. 19-36. Oxford University Press, Oxford, 1990.

Osterhout, L., and Holcomb, P. J. Event-related potentials and language comprehension. In *Electrophysiology of Mind: Event-Related Potentials and Cognition.* M. D. Rugg, and M. G. H. Coles (Eds), pp. 171-215. Oxford University Press, Oxford, 1995.

Paller, K. A., and Kutas, M. Brain potentials during retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience* **4**, 375-391, 1992.

Paller, K. A., Kutas, M., and Mayes, A. R. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology* **67**, 360-371, 1987.

Paller, K. A., Kutas, M., and McIsaac, H. K. Monitoring conscious recollection via the electrical activity of the brain. *Psychological Science* **6**, 107-111, 1995.

Paller, K. A., McCarthy, G., and Wood, C. C. ERPs predictive of subsequent recall and recognition performance. *Biological Psychology* **26**, 269-276, 1988.

Parkin, A. J., Reid, T. K., and Russo, R. On the differential nature of implicit and explicit memory. *Memory and Cognition* **18**, 507-514, 1990.

Parkin, A. J., and Walter, B. M. Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging* **7**, 290-298, 1992.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology* **72**, 184-187, 1989.

Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., and Echallier, J. F. Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology* **66**, 75-81, 1987.

Picton, D. W., and Stuss, D. T. The component structure of the human event-related potentials. In *Progress in Brain Research.* H. H. Kornhuber, and L. Deecke (Eds), pp. 17-49. Elsevier/North Holland Biomedical Press, Amsterdam, 1980.

Picton, T. W. The recording and measurement of evoked potentials. In *A Textbook of Clinical Neurophysiology.* A. M. Halliday, S. R. Butler, and R. Paul (Eds), pp. 23-40. Wiley, New York, 1987.

Picton, T. W., Lins, O. G., and Scherg, M. The recording and analysis of event-related potentials. In *Handbook of Neuropsychology, Volume 10.* F. Boller, and J. Grafman (Eds), pp. 429-499. Elsevier, Amsterdam, 1995.

Potter, D. D., Pickles, C. D., Roberts, R. C., and Rugg, M. D. The effects of scopolamine on event-related potentials in a continuous recognition memory task. *Psychophysiology* **29**, 29-37, 1992.

Pritchard, W. Psychophysiology of P300. *Psychological Bulletin* **89**, 506-540, 1981.

Rajaram, S. Remembering and knowing: Two means of access to the personal past. *Memory and Cognition* **21**, 89-102, 1993.

Ratcliff, R., and McKoon, G. More on the distinction between episodic and semantic memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 312-313, 1986.

Reed, A. V. Speed-accuracy tradeoff in recognition memory. *Science* **181**, 574-576, 1973.

Richardson-Klavehn, A., and Bjork, R. A. Measures of memory. *Annual Review of Psychology* **39**, 475-543, 1988.

Roediger, H. L. The problem of differing false-alarm rates for the process dissociation procedure: Comment on Verfaellie and Treadwell (1993). *Neuropsychology* **8**, 284-288, 1994.

Roediger, H. L., Weldon, M. S., and Challis, B. H. Explaining dissociations between implicit and explicit measures of retention: A processing account. In *Varieties of Memory and Consciousness.* H. L. Roediger, and F. I. M. Craik (Eds), pp. 3-41. Erlbaum, Hillsdale, New Jersey, 1989.

Roland, P.E., Kawashima, R., Gulyas, B., and O'Sullivan, B. Positron emission tomography in cognitive neuroscience: Methodological constraints, strategies, and examples from learning and memory. In *The Cognitive Neurosciences*, M.S. Gazzaniga (Ed), pp. 781-788, MIT Press, Cambridge, Massachusetts, 1994.

Ruchkin, D. E., and Sutton, S. Emitted P300s and temporal uncertainty. *Electroencephalography and Clinical Neurophysiology* **45**, 268-277, 1978.

Rugg, M. D. The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology* **22(6)**, 642-647, 1985.

Rugg, M. D. Event-related brain potentials dissociate repetition effects of high and low frequency words. *Memory and Cognition* **18**, 367-379, 1990.

Rugg, M. D. Event-related potentials in clinical neuropsychology. In *A Handbook of Neuropsychological Assessment.* J. R. Crawford, D. M. Parker, and W. W. McKinlay (Eds), pp. 393-411. Erlbaum, Hove, 1992.

Rugg, M. D. ERP Studies of memory. In *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition.* M. D. Rugg, and M. G. H. Coles (Eds), pp. 113-127. Oxford University Press, London, 1994.

Rugg, M. D. Cognitive event-related potentials: Intracranial and lesion studies. In *Handbook of Neuropsychology, Volume 10.* J. C. Baron, and J. Grafman (Eds), pp. 165-186. Elsevier, Amsterdam, 1995.

Rugg, M. D., Allan, K. A., and Wilding, E. L. An investigation of the processes contributing to performance on recognition memory exclusion tasks. *Paper presented at Memory Disorders Society Meeting, Cambridge, August, 1995.*

Rugg, M. D., and Barrett, S. E. Event-related potentials and the interaction between orthographic and phonological information in a rhyme-judgment task. *Brain and Language* **32**, 336-361, 1987.

Rugg, M. D., Brovedani, P., and Doyle, M. C. Modulation of event-related potentials by word repetition in a task with inconsistent mapping between repetition and response. *Electroencephalography and Clinical Neurophysiology* **84**, 521-531, 1992.

Rugg, M. D., and Coles, M. G. H. The ERP and cognitive psychology: Conceptual issues. In *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition.* M. D. Rugg, and M. G. H. Coles (Eds), pp. 27-39. Oxford University Press, Oxford, 1994.

Rugg, M. D., Cox, C. J. C., Doyle, M. C., and Wells, T. Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia* **33**, 471-484, 1995.

Rugg, M. D., and Doyle, M. C. Event-related potentials and recognition memory for low- and high- frequency words. *Journal of Cognitive Neuroscience* **4**, 69-79, 1992.

Rugg, M. D., and Doyle, M. C. Event-related potentials and stimulus repetition in direct and indirect tests of memory. In *Cognitive Electrophysiology.* H. J. Heinze, T. Munte, and G. R. Mangun (Eds), pp. 124-148. Birkhauser, Boston, 1994.

Rugg, M. D., Doyle, M. C., and Melan, C. An event-related potential study of the effects of intra- and inter- modal word repetition. *Language and Cognitive Processes* **8**, 357-378, 1993.

Rugg, M. D., Fletcher, P. C., Frith, C. D., Frackowiack, R. S. J., and Dolan, R. J. Differential response of the prefrontal cortex in successful and unsuccessful memory retrieval. *submitted.*

Rugg, M. D., Furda, J., and Lorist, M. The effects of task on the modulation of event-related potentials by word repetition. *Psychophysiology* **25**, 55-63, 1988.

Rugg, M. D., and Nagy, M. E. Lexical contribution to non-word repetition effects: Evidence from event-related potentials. *Memory and Cognition* **15**, 473-481, 1987.

Rugg, M. D., and Nagy, M. E. Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology* **72**, 395-406, 1989.

Rugg, M. D., Roberts, R. C., Potter, D. D., Pickles, C. D., and Nagy, M. E. Event-related potentials related to recognition memory. Effects of unilateral temporal lobectomy and temporal lobe epilepsy. *Brain* **114**, 2313-2332, 1991.

Russo, R., and Andrade, J. The directed forgetting effect in word fragment completion: An application of the process dissociation procedure. *Quarterly Journal of Experimental Psychology* **48A(2)**, 405-423, 1995.

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., and Lindsley, D. B. Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology* **17**, 568-576, 1980.

Scarborough, D. L., Gerard, L., and Cortese, C. Accessing lexical memory: The transfer of word recognition effects across task and modality. *Memory and Cognition* **7**, 3-12, 1979.

Schacter, D. L. Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 501-518, 1987.

Schacter, D. L., Bowers, J. S., and Booker, J. Intention, awareness, and implicit memory: The retrieval intentionality criterion. In *Implicit Memory: Theoretical Issues.* S. Lewandowsky, J. Dunn, and K. Kirsner (Eds), pp. 47-65. Erlbaum, Hillsdale, New Jersey, 1989.

Schacter, D. L., Harbluk, J. L., and McLachlan, D. R. Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behaviour* **23**, 593-611, 1984.

Scherg, M. Fundamentals of dipole source potential analysis. In *Auditory Evoked Magnetic fields and Electric Potentials: Advances in Audiology.* F. Grandori, M. Hoke, and G. L. Romani (Eds), pp. 40-69. Karger, Basel, 1990.

Schloerscheidt, A. M., Rugg, M. D., Doyle, M. C., Cox, C. J. C., and Patching, G. R. Recognition memory with and without retrieval of the encoding episode: A study with event-related brain potentials. *Paper presented at conference of the European Neuroscience Association, Amsterdam, 3-7 September, 1995.*

Scoville, W. B., and Milner, B. Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry* **20**, 11-21, 1957.

Shallice, T., and Warrington, E. K. Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology* **22**, 261-273, 1970.

Shimamura, A. P. Problems with the finding of stochastic independence as evidence for multiple memory systems. *Bulletin of the Psychonomic Society* **23**, 506-508, 1985.

Shimamura, A. P. Priming effects in amnesia: Evidence for a dissociable memory function. *Quarterly Journal of Experimental Psychology* **38A**, 619-644, 1986.

Shimamura, A. P. Memory and frontal lobe function. In *The Cognitive Neurosciences.* M. S. Gazzaniga (Ed), pp. 803-813. MIT Press, Cambridge, Massachusetts, 1994.

Shimamura, A. P., and Squire, L. R. Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General* **113**, 556-570, 1984.

Shimamura, A. P., and Squire, L. R. A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 464-473, 1987.

Smith, M. E. Neurophysiological manifestations of recollective experience during recognition memory judgements. *Journal of Cognitive Neuroscience* **5**, 1-13, 1993.

Smith, M. E., and Guster, K. Decomposition of recognition memory event-related potentials yields target, repetition, and retrieval effects. *Electroencephalography and Clinical Neurophysiology* **86**, 335-343, 1993.

Smith, M. E., and Halgren, E. Dissociation of recognition memory components following temporal lobe lesions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 50-60, 1989.

Smith, M. E., Stapleton, J. M., and Halgren, E. Human medial temporal-lobe potentials evoked in memory and language tasks. *Electroencephalography and Clinical Neurophysiology* **63**, 145-159, 1986.

Snodgrass, J. G., and Corwin, J. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General* **117**, 34-50, 1988.

Squire, L. R. The neuropsychology of human memory. *Annual Review of Neuroscience* **5**, 241-273, 1982a.

Squire, L. R. Comparisons between forms of amnesia: Some deficits are unique to Korsakoff's syndrome. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **8**, 560-571, 1982b.

Squire, L. R. Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience* **4**, 232-243, 1992.

Squire, L. R. Declarative and non-declarative memory: Multiple brain systems supporting learning and memory. In *Memory Systems 1994.* E. Tulving, and D. L. Schachter (Eds), pp. 203-231. MIT Press, Cambridge, Massachusetts, 1994.

Squire, L. R., Knowlton, B., and Musen, G. The structure and organization of memory. *Annual Review of Psychology* **44**, 453-495, 1993.

Squire, L. R., and Knowlton, B. J. Memory, hippocampus, and brain systems. In *The Cognitive Neurosciences.* M. S. Gazzaniga (Ed), pp. 825-837. MIT Press, Cambridge, Massachusetts, 1994.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., and Raichle, M. E. Activation of the hippocampus in normal humans: A functional anatomical study of memory. *Proceedings of the National Academy of Sciences* **89**, 1837-1841, 1992.

Squire, L. R., Shimamura, A. P., and Graf, P. Independence of recognition memory and priming effects: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11**, 37-44, 1985.

Squire, L. R., and Zola-Morgan, M. The medial temporal lobe memory system. *Science* **253**, 1380-1386, 1991.

Squire, L. R., and Zola-Morgan, S. Memory: Brain systems and behavior. *Trends in Neuroscience* **11**, 170-175, 1988.

Stapleton, J. M., Halgren, E., and Moreno, K. A. Endogenous potentials after anterior temporal lobectomy. *Electroencephalography and Clinical Neurophysiology* **67**, 44-52, 1987.

Sternberg, S. High-speed scanning in human memory. *Science* **153**, 652-654, 1966.

Stuss, D. T., Eskes, G. A., and Foster, J. K. Experimental neuropsychological studies of frontal lobe functions. In *Handbook of Neuropsychology, Volume 9.* F. Boller, and J. Grafman (Eds), pp. 149-183. Elsevier, Amsterdam, 1994.

Sutton, S., Braren, M., Zubin, J., and John, E. R. Evoked potential correlates of stimulus uncertainty. *Science* **150**, 1187-1188, 1965.

Toth, J. P., Reingold, E. M., and Jacoby, L. L. A response to Graf and Komatsu's critique of the process dissociation procedure: When is caution necessary ? *European Journal of Cognitive Psychology* **7**, 113-130, 1995.

Tulving, E. Episodic and semantic memory. In *Organization of Memory.* E. Tulving, and W. Donaldson (Eds), pp. 381-403. Academic Press, New York, 1972.

Tulving, E. *Elements of Episodic Memory*, Oxford University Press, Oxford, 1983.

Tulving, E. Memory and Consciousness. *Canadian Psychologist* **26**, 1-12, 1985.

Tulving, E. What kind of hypothesis is the distinction between episodic and semantic memory ? *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 307-311, 1986.

Tulving, E. Human memory. In *Memory Concepts - 1993. Basic and Clinical Aspects.* P. Andersen, O. Hvalby, O. Paulsen, and B. Hokfelt (Eds), pp. 27-45. Elsevier, New York, 1993a.

Tulving, E. What is episodic memory ? *Current Directions in Psychological Science* **2**, 67-70, 1993b.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., and Houle, S. Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Science* **91**, 2016-2020, 1994a.

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I. M., Habib, R., and Houle, S. Neuroanatomical correlates of retrieval in episodic memory: auditory sentence recognition. *Proceedings of the National Academy of Science* **91**, 2012-2015, 1994b.

Tulving, E., Schacter, D. L., and Stark, H. A. Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **8**, 352-373, 1982.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., and McIsaac, H. Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience* **3**, 129-150, 1991.

Verfaellie, M. A re-examination of recognition memory in amnesia: Reply to Roediger and McDermott. *Neuropsychology* **8**, 289-292, 1994.

Verfaellie, M., and Treadwell, J. R. Status of recognition memory in amnesia. *Neuropsychology* **7**, 5-13, 1993.

Warrington, E. K., and Weiskrantz, L. New method of testing long-term retention with special reference to amnesic patients. *Nature* **217**, 972-974, 1968.

Warrington, E. K., and Weiskrantz, L. Amnesic syndrome: Consolidation or retrieval ? *Nature* **228**, 628-630, 1970.

Warrington, E. K., and Weiskrantz, L. The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia* **12**, 419-428, 1974.

Waugh, N., and Norman, D. A. Primary memory. *Psychological Review* **72**, 89-104, 1965.

Weldon, M. S., and Roediger, H. L. Altering retrieval demands reverses the picture superiority effect. *Memory and Cognition* **15**, 269-280, 1987.

Whittlesea, B. W. A. Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **19**, 1235-1253, 1993.

Whittlesea, B. W. A., Jacoby, L. L., and Girard, K. A. Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language* **29**, 716-732, 1990.

Wickelgren, W. A. Sparing of short-term memory in an amnesic patient: Implications for strength theory of memory. *Neuropsychologia* **6**, 235-244, 1968.

Wickelgren, W. A. *Dynamics of Retrieval*, Academic Press, New York, 1975.

Wilding, E. L., Doyle, M. C., and Rugg, M. D. Recognition memory with and without retrieval of study context: An event-related potential study. *Neuropsychologia* **33**, 743-767, 1995.

Winer, B. J. *Statistical Principles in Experimental Design, 2nd Edition*, McGraw-Hill, New York, 1971.

Wood, C. C. Application of dipole localization methods to source identification of human evoked potentials. *Annals of the New York Academy of Sciences* **388**, 139-155, 1982.

Wood, C. C. Generators of event-related potentials. In *A Textbook of Clinical Neurophysiology.* A. M. Halliday, S. M. Butler, and R. Paul (Eds), pp. 535-567. Wiley, New York, 1987.

Wood, C. C., McCarthy, G., Allison, T., Goff, W. R., Williamson, P. D., and Spencer, D. D. Endogenous event-related potentials following temporal lobe excisions in humans. *Society for Neuroscience Abstracts* **8**, 976, 1982.

Yonelinas, A. P. Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20**, 1341-1354, 1994.

Yonelinas, A. P., and Jacoby, L. L. Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology* **48**, 516-534, 1994.

Yonelinas, A. P., and Jacoby, L. L. The relationship between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language* **34**, 622-643, 1995.

Zola-Morgan, S., and Squire, L. R. The neuroanatomy of amnesia. *Annual Review of Neuroscience* **16**, 547-563, 1993.

Appendix 1

Mean amplitude measures for sites included in the critical ERP analyses in experiments 1 to 6.

Appendix 1.1 Mean amplitude (μV) of the ERPs evoked by the Visual Hit/Hit (V Hit/Hit), Auditory Hit/Hit (A Hit/Hit), Auditory Hit/Miss (A Hit/Miss) and Correct Rejection (CR) response categories in experiment 1. The mean amplitudes displayed are for the 300-500 and 500-900 msec epochs. Fz, Cz, and Pz signify midline frontal, central, and parietal sites. LF, RF, LT, RT, LP, RP, T5, T6, O1, O2 signify left and right, frontal, anterior temporal, parietal, posterior temporal and occipital sites. Note that all subsequent appendices showing mean amplitude measures do so for the electrode sites listed here.

	Fz	Cz	Pz	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
300-500													
V Hit/Hit:	-0.02	-2.21	1.29	-0.98	-1.49	0.82	-0.10	0.56	1.21	-0.63	0.99	0.50	0.92
A Hit/Hit:	-0.36	-1.58	1.02	-1.00	-1.28	0.67	0.25	0.19	0.93	-1.17	0.76	0.22	0.39
A Hit/Miss:	-1.34	-2.48	0.03	-2.23	-2.81	-0.89	-1.72	-0.68	0.27	-1.66	0.13	-0.22	-0.24
CR:	-1.27	-2.67	0.34	-1.99	-2.38	-0.20	-0.73	-0.61	0.21	-1.53	0.38	-0.18	0.19
500-900													
V Hit/Hit:	3.07	5.16	6.50	2.27	4.19	6.21	4.05	3.62	3.83	3.73	3.40	1.84	3.20
A Hit/Hit:	2.98	6.00	6.65	2.32	4.73	5.97	3.98	3.54	4.06	3.57	3.53	1.81	2.99
A Hit/Miss:	0.36	4.02	4.76	0.28	2.29	3.95	1.68	2.31	2.11	1.90	2.03	0.71	1.88
CR:	0.57	3.29	5.13	0.27	2.28	3.68	2.17	2.44	2.12	2.45	2.57	0.93	2.36

Appendix 1.2 Mean amplitude (μV) of the ERPs evoked by the Auditory Hit/Hit (A Hit/Hit), Visual Hit/Hit (V Hit/Hit), collapsed Hit/Miss and Correct Rejection (CR) response categories in experiment 2. The mean amplitudes are shown for the 400-800, 800-1100, and 1100-1400 msec epochs.

	Fz	Cz	Pz	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
400-800													
AHit/Hit:	-10.10	-10.29	-5.52	-10.07	-9.01	-3.39	-1.90	-1.42	-8.48	-8.86	-3.15	-2.38	-1.65
VHit/Hit:	-10.38	-10.81	-6.35	-9.86	-8.85	-4.09	-2.68	-1.65	-8.10	-8.42	-3.36	-2.49	-2.18
Hit/Miss:	-10.17	-10.00	-5.76	-10.20	-8.46	-3.62	-1.85	-1.71	-8.02	-7.67	-3.07	-2.16	-1.81
CR:	-11.76	-12.18	-7.50	-11.38	-10.14	-4.98	-3.01	-2.32	-9.61	-9.42	-4.07	-2.88	-2.81
800-1100													
AHit/Hit:	-6.66	-4.11	-0.04	-7.65	-4.75	2.99	2.81	3.13	-5.70	-4.47	2.33	1.27	2.31
VHit/Hit:	-7.30	-5.13	-0.41	-7.96	-4.81	2.57	2.20	3.54	-5.93	-4.75	1.96	1.57	2.37
Hit/Miss:	-8.31	-5.97	-1.87	-8.70	-5.52	1.37	1.94	2.11	-6.77	-5.21	0.59	0.04	1.26
CR:	-8.96	-6.36	-1.72	-9.45	-6.33	0.77	1.15	2.12	-7.90	-5.93	0.63	0.35	1.00
1100-1400													
AHit/Hit	-4.30	-3.49	-2.97	-5.57	-3.87	0.84	1.17	0.11	-2.74	-1.87	0.49	-0.01	-0.30
VHit/Hit	-4.24	-2.40	-0.74	-4.52	-2.20	2.88	2.21	2.44	-1.99	-0.00	1.87	1.50	1.48
Hit/Miss	-5.31	-3.91	-3.05	-6.13	-3.87	0.67	0.87	0.38	-3.27	-2.09	0.28	0.13	-0.05
CR	-4.82	-2.92	-1.46	-5.40	-3.69	0.72	0.65	0.83	-3.82	-2.36	0.20	0.14	-0.13

Appendix 1.3 Mean amplitude (μV) of the ERPs evoked by the Hit/Hit, Hit/Miss and Correct Rejection (CR) response categories in experiment 3. The mean amplitudes are shown for the 500-800, 800-1100, and 1100-1400 msec epochs.

	FZ	CZ	PZ	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
500-800													
Hit/Hit:	3.72	6.69	7.56	3.80	6.07	7.61	3.99	2.93	4.06	3.32	4.70	2.16	1.77
Hit/Miss:	1.15	5.48	8.00	1.93	5.00	6.84	3.52	3.28	1.85	2.29	4.44	1.89	2.43
CR:	0.15	4.22	6.70	1.18	3.60	4.95	2.45	2.16	1.08	1.55	3.44	1.35	1.56
800-1100													
Hit/Hit:	6.07	7.36	4.42	5.20	6.82	7.07	4.26	2.62	7.19	6.52	5.88	3.49	1.75
Hit/Miss:	3.48	6.32	4.71	3.63	5.69	6.04	3.38	2.50	4.89	5.06	4.87	2.67	2.05
CR:	4.15	7.09	5.84	3.85	5.46	4.98	2.83	2.51	5.54	5.54	4.87	2.50	1.97
1100-1400													
Hit/Hit	5.49	4.05	-0.56	3.41	2.97	1.60	0.37	-0.76	8.21	5.45	2.50	2.01	-0.04
Hit/Miss	3.74	3.96	0.67	2.45	2.76	2.06	0.74	-0.08	6.21	4.71	2.73	1.91	0.33
CR	3.02	3.96	1.94	2.30	2.43	1.22	0.17	0.23	4.97	4.07	2.60	1.68	0.39

Appendix 1.4 Mean amplitude (μV) of the ERPs evoked by the Hit/Hit, Hit/Miss and Correct Rejection (CR) response categories in experiment 4. The mean amplitudes are shown for the 500-800, 800-1100, and 1100-1400 msec epochs.

	FZ	CZ	PZ	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
500-800													
Hit/Hit:	2.59	6.20	8.08	1.85	5.21	6.73	3.71	5.07	2.96	3.57	4.17	1.23	4.17
Hit/Miss:	1.05	4.46	6.27	0.75	3.34	4.85	2.49	3.73	1.53	2.26	2.29	-0.18	2.97
CR:	-0.87	2.96	5.89	-0.74	2.16	3.34	1.23	3.23	-0.10	1.33	1.83	-0.35	2.75
800-1100													
Hit/Hit:	6.85	9.04	6.28	6.02	7.71	6.35	3.93	2.91	8.75	8.30	5.53	2.26	1.66
Hit/Miss:	4.19	7.08	4.05	4.26	5.39	3.98	2.31	1.47	6.25	6.04	2.88	0.34	0.49
CR:	3.46	7.08	5.37	3.93	5.17	3.40	1.55	1.62	4.92	5.65	3.39	1.04	1.01
1100-1400													
Hit/Hit	6.92	6.34	2.10	5.02	5.29	2.51	0.95	-0.65	9.82	8.01	3.43	1.03	-1.32
Hit/Miss	4.89	5.45	0.96	3.81	3.72	1.40	0.49	-0.98	7.69	6.25	1.59	-0.28	-1.62
CR	3.64	5.46	2.96	3.77	4.19	1.84	0.65	0.22	5.15	5.23	2.38	0.43	-0.38

Appendix 1.5 Mean amplitude (μV) of the ERPs evoked by the Non-Target hit (NT Hit), Target Miss (T Miss) and Correct Rejection (CR) response categories in experiment 5. The mean amplitudes are shown for the 500-800, 800-1100, and 1100-1400 msec epochs.

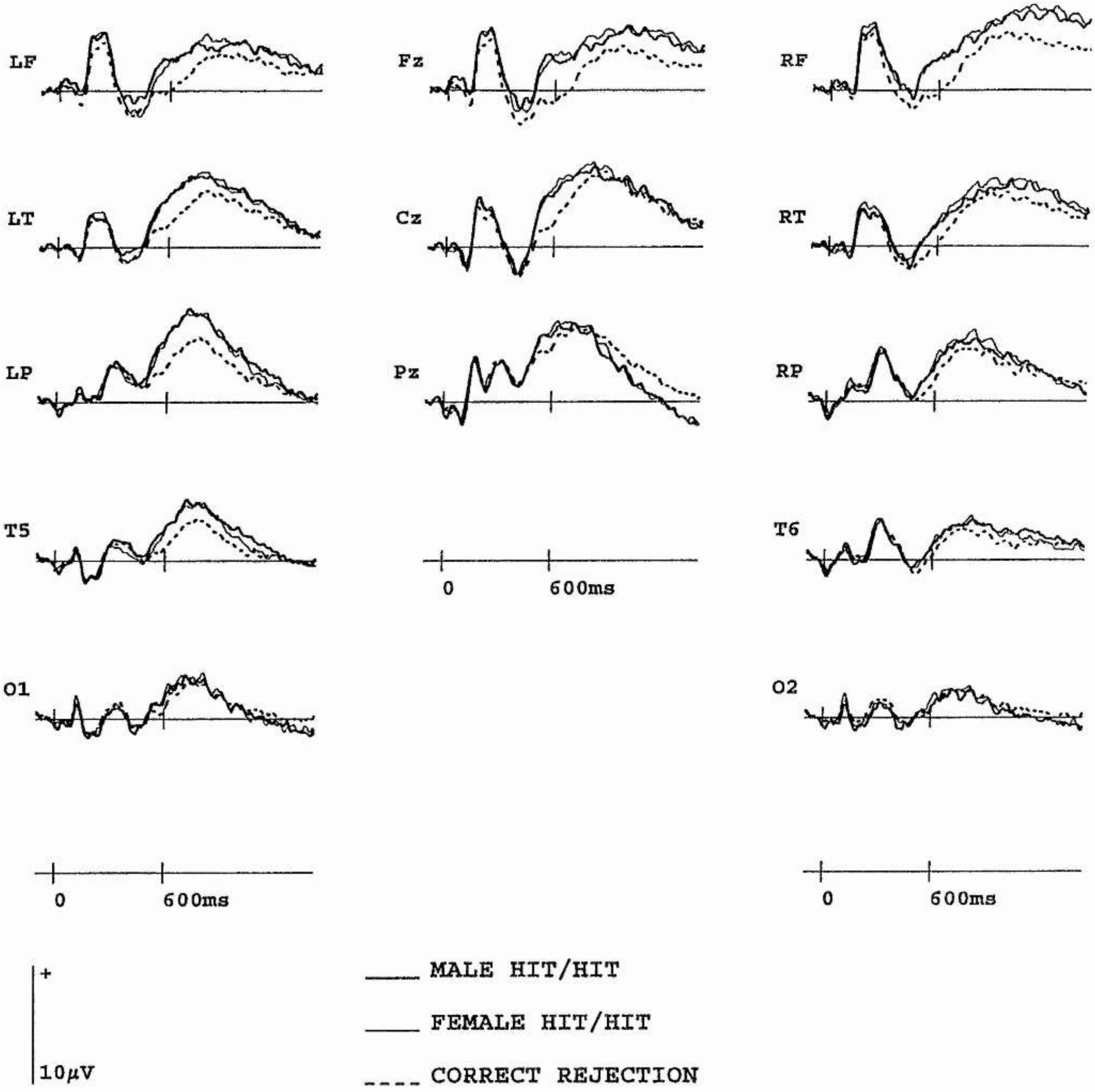
	FZ	CZ	PZ	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
500-800													
NT Hit:	-2.44	0.96	3.56	-1.63	1.06	2.25	0.54	0.80	-2.36	-0.86	0.21	-0.99	0.72
T Miss:	-3.78	0.12	2.97	-2.38	0.07	1.49	-0.35	0.28	-2.15	-0.75	0.25	-0.29	0.63
CR:	-3.26	-0.20	2.85	-2.21	0.09	0.96	-0.60	0.04	-2.17	-0.89	0.07	-1.00	0.10
800-1100													
NT Hit:	-0.45	0.62	0.73	-0.04	1.51	1.60	0.43	-0.17	1.17	1.45	0.88	-0.34	-0.45
T Miss:	-0.49	1.54	1.29	0.01	1.25	1.40	-0.33	-0.10	1.57	2.31	1.49	0.81	0.16
CR:	0.05	2.02	2.32	0.23	1.49	1.14	-0.35	0.18	2.18	2.73	1.87	0.70	0.27
1100-1400													
NT Hit:	-0.57	-1.47	-1.71	-0.74	-0.43	-0.22	-0.51	-0.80	1.41	0.91	0.29	-0.03	-0.89
T Miss:	-0.40	0.21	-0.04	-0.49	-0.03	0.82	-0.35	0.04	1.72	2.32	1.62	1.40	0.28
CR:	-0.01	0.74	1.02	-0.29	-0.04	0.46	-0.53	0.26	2.09	2.08	2.00	1.45	0.58

Appendix 1.6 Mean amplitude (μV) of the ERPs evoked by the Spoken Hit/Hit (Sp Hit/Hit), Heard Hit/Hit (H Hit/Hit) and Correct Rejection (CR) response categories in experiment 6. The mean amplitudes are shown for the 500-800, 800-1100, and 1100-1400 msec epochs.

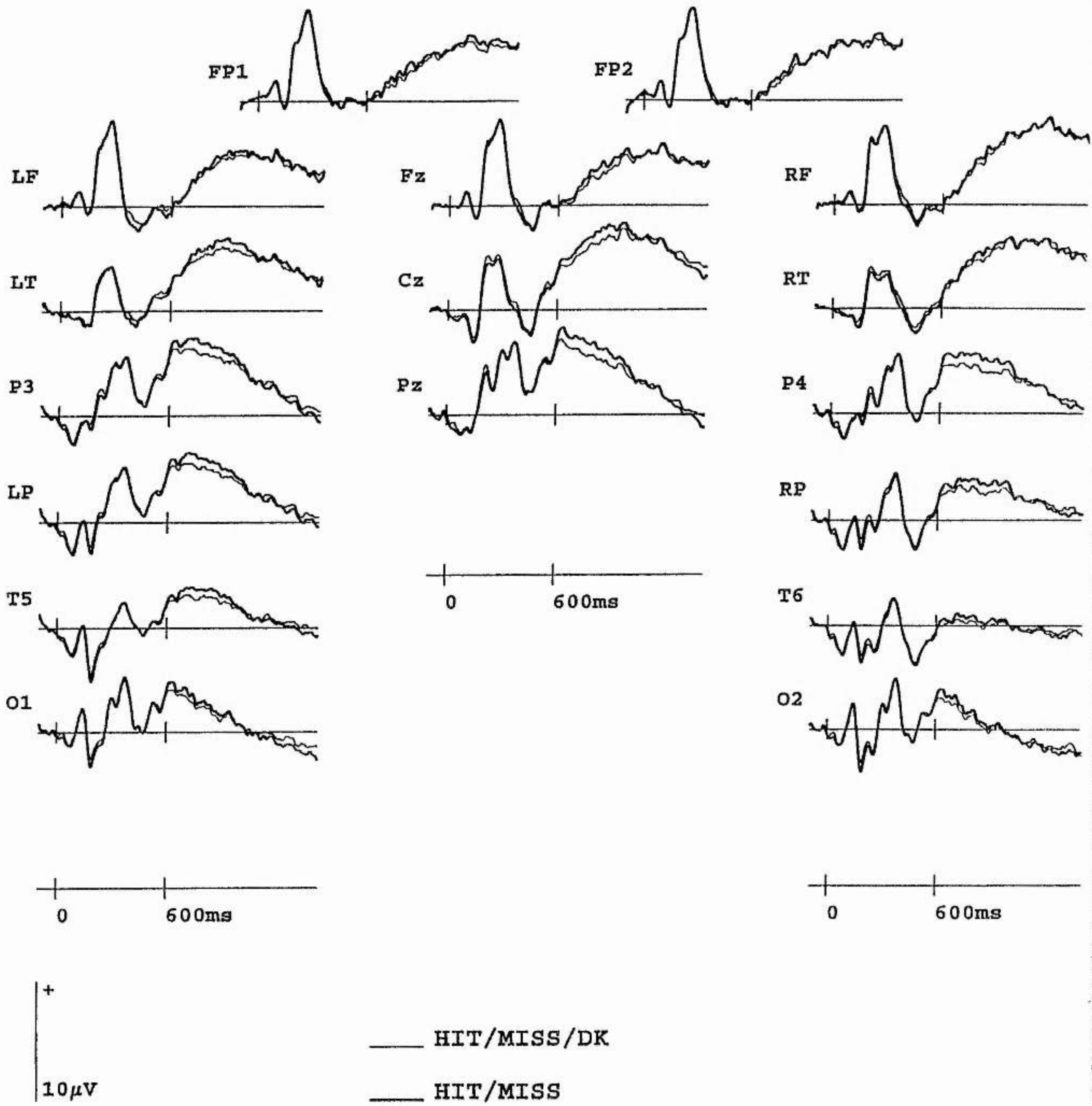
	FZ	CZ	PZ	LF	LT	LP	T5	O1	RF	RT	RP	T6	O2
500-800													
Sp Hit/Hit:	6.20	7.89	6.13	5.71	7.47	6.22	3.37	1.85	4.61	4.31	2.64	0.01	0.62
H Hit/Hit:	4.11	5.89	4.57	3.80	5.34	3.74	1.39	0.40	2.88	2.74	1.17	-0.98	-0.03
CR:	1.50	3.04	1.87	1.65	2.96	0.77	-0.80	-0.99	1.09	1.31	-0.51	-1.69	-1.38
800-1100													
Sp Hit/Hit:	4.67	3.78	-0.43	4.57	4.69	2.57	1.01	-1.42	5.36	4.17	0.29	-1.68	-2.45
H Hit/Hit:	4.59	4.13	-0.19	4.58	4.46	1.78	0.57	-1.77	5.11	4.01	0.07	-1.53	-2.04
CR:	2.04	3.64	0.46	2.51	2.99	0.21	-0.98	-1.63	3.13	3.27	-0.52	-1.52	-1.91
1100-1400													
Sp Hit/Hit:	2.16	-1.24	-6.13	2.17	1.56	-1.45	-1.62	-3.95	3.79	2.00	-2.37	-3.10	-4.56
H Hit/Hit:	2.74	-0.01	-4.64	3.20	1.81	-1.32	-1.44	-3.80	3.97	2.31	-2.26	-3.03	-3.93
CR:	1.58	1.53	-1.64	1.81	1.24	-1.30	-1.79	-2.03	2.44	2.14	-1.18	-1.78	-1.96

Appendix 2

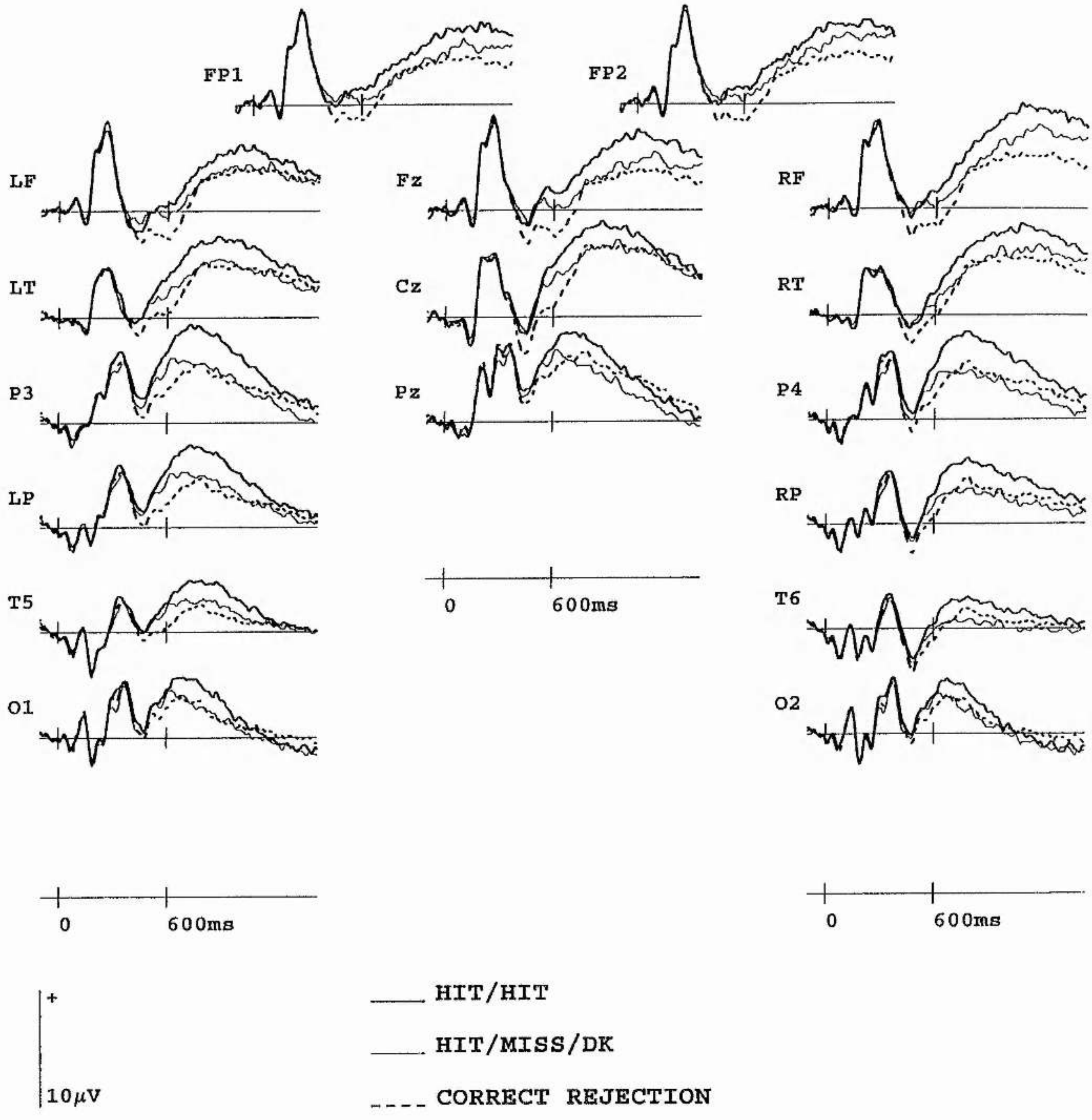
Additional ERP figures.



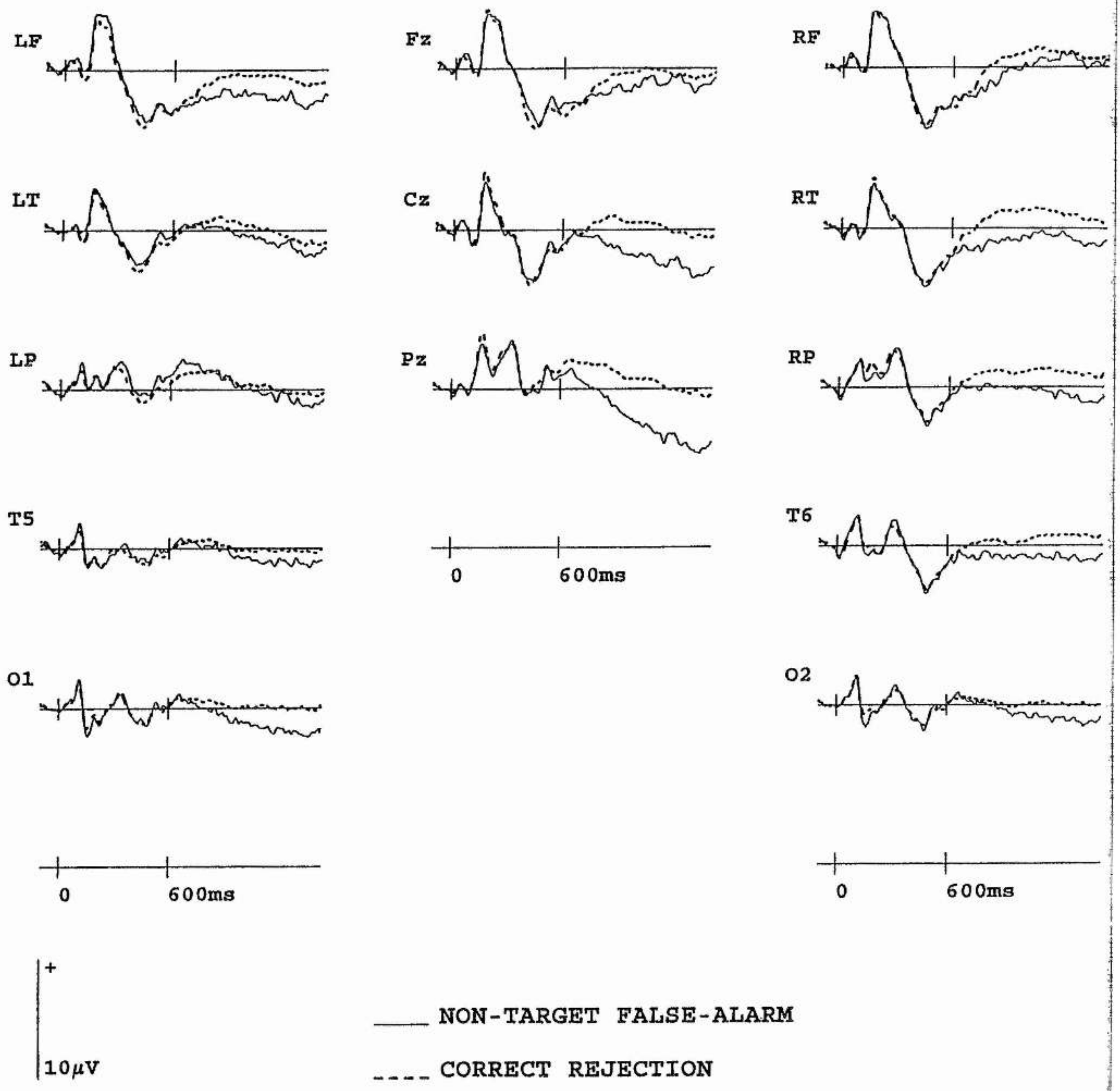
Appendix 2.1 Grand average ERPs associated with the male hit/hit, female hit/hit, and correct rejection response categories in experiment 3. Fz, Cz, and Pz signify midline frontal, central, and parietal sites. LF, RF, LT, RT, LP, RP, T5, T6, O1, O2 signify left and right frontal, anterior temporal, parietal, posterior temporal and occipital sites.



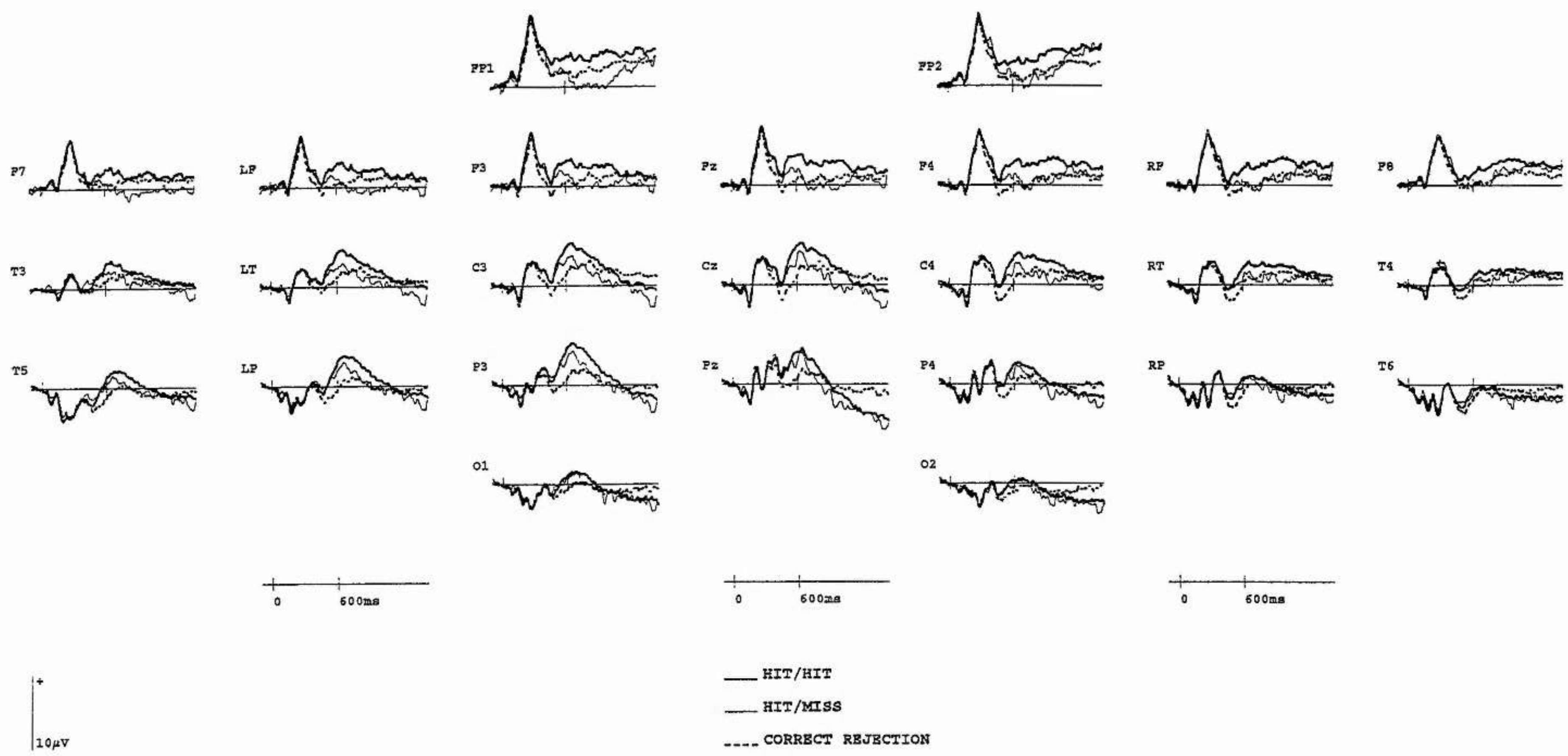
Appendix 2.2 Grand average ERPs associated with the hit/miss response category and the response category formed by collapsing across the hit/miss and don't know response categories (hit/miss/dk) in experiment 4. Electrode sites as for appendix 2.1, with four additional sites: FP1, FP2, P3, and P4, which correspond to left and right, prefrontal and superior parietal sites.



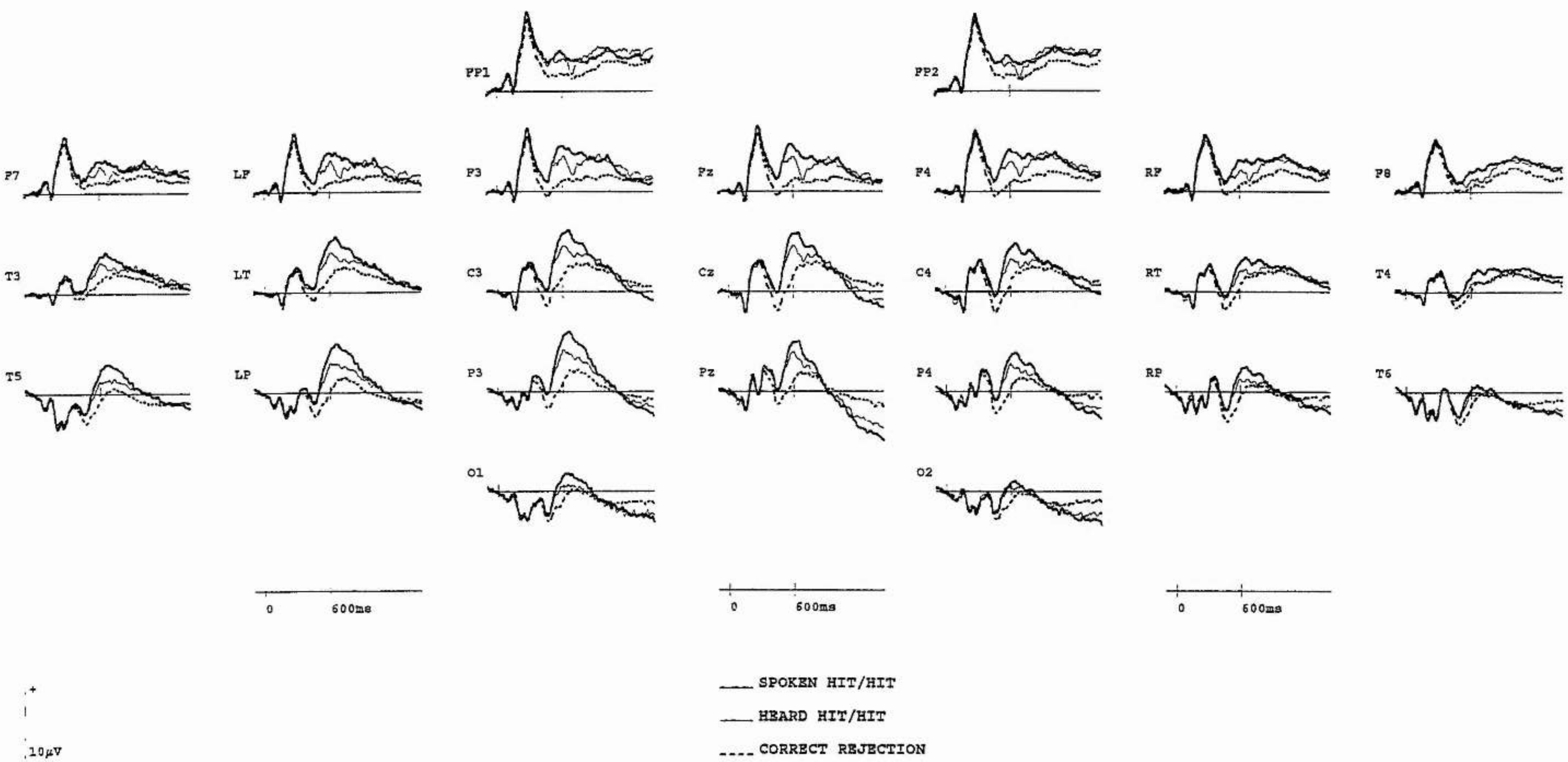
Appendix 2.3 Grand average ERPs associated with the hit/hit, hit/miss/dk, and correct rejection response categories in experiment 3. Electrode sites as for appendix 2.2.



Appendix 2.4 Grand average ERPs associated with the non-target false alarm and correct rejection response categories in experiment 5. Electrode sites as for appendix 2.1.



Appendix 2.5 Grand average ERPs associated with the collapsed hit/hit, hit/miss and correct rejection response categories in experiment 6. Electrode sites as for appendix 2.2, with the following additional sites: F7, F8, F3, F4, T3, T4, C3, and C4, which correspond to left and right inferior frontal, superior frontal, inferior temporal, and superior temporal electrode locations.



Appendix 2.6 Grand average ERPs associated with the spoken hit/hit, heard hit/hit and correct rejection response categories in experiment 6.

Electrode sites as for appendix 2.5.

Appendix 3

Lists of word and non-word stimuli from which the experimental stimuli used in experiments 1 to 6 were drawn.

Appendix 3.1 Pool of words from which experimental stimuli were drawn

accordion	caper	digital	frown	jewel	monkey	pulpit	sermon
acid	cargo	disciple	gasket	juice	monster	puppy	shark
adapter	casualty	discount	gasp	kindle	morsel	purge	shawl
adjacent	chant	dissolve	gauge	kneel	mortal	quack	sheep
adultery	chap	distrust	gaunt	lace	moth	quaint	shin
adviser	chaplain	divert	gauntlet	lagoon	mound	quart	shrapnel
alcohol	check	dogma	germ	lance	mouse	raft	shrimp
alphabet	cheese	donor	germane	lantern	muck	rafter	shrink
antler	chimney	dove	ghetto	lapse	mule	raisin	shuffle
anvil	choke	dragon	gloss	lard	mutant	ramble	sickness
apathy	chord	drawl	glossary	larder	mute	ramp	sigh
apex	cigar	duck	glum	latch	neglect	rampage	silk
applause	clam	dung	glut	layout	newt	rash	skewer
appoint	claw	dungeon	goad	leap	nudge	rat	skillet
apricot	cliff	durable	gossip	leash	oracle	rave	skip
attic	cling	easel	granary	legacy	otter	raven	skull
awkward	cloak	elder	grape	ligament	padlock	recruit	skunk
bachelor	clog	engage	grave	lilac	pagan	remedy	sledge
baggage	coarse	enzyme	gravy	limb	parapet	rental	sleet
bandage	coil	escapade	grim	lion	parasite	retina	slim
bang	comet	ethnic	grimace	listener	parcel	retinue	slip
bangle	comply	evoke	groan	lobster	partial	revere	slog
bargain	contour	explore	grope	loom	passport	reversal	slot
barge	copious	fantasy	growl	lunar	pastel	rinse	slumber
bead	cordon	fatigue	gypsy	lunatic	patio	roam	smog
beaker	cough	feast	halt	lunge	patriot	roar	snail
beast	cove	feat	halter	lure	peach	robin	snake
bison	coward	feather	hammer	lurid	pelvis	robot	snare
bite	crab	feeble	harp	majestic	pendant	rocky	sniff
blade	crank	fern	harsh	marital	pension	rogue	snore
blast	crate	fetch	hatch	marsh	pert	rubber	snorkel
blaze	crater	feud	hatchet	mask	pickle	rumble	solar
blazer	crawl	figment	haven	massage	pivot	rusk	solvent
bleat	crescent	flange	hawk	mate	plank	rust	sonata
bloom	crimp	flank	heat	matrix	pliers	saint	soprano
brandish	crimson	flask	heathen	matron	plum	sapling	space
breech	crook	flea	helm	maze	plume	saucepan	spade
bribe	crow	flock	helmet	melon	podium	scalp	spangle
bronze	crucifix	flood	hive	mesh	poise	scan	sparse
bruise	crush	florist	horse	meter	poison	scandal	spasm
budge	crypt	flush	host	mild	pony	scar	spear
bulb	cryptic	flute	hound	mill	pouch	scare	spice
bull	cube	foal	hump	mineral	pram	scarf	spider
burglar	culprit	force	hurt	mink	prank	scream	spin
burly	cupboard	fortress	instinct	mint	prawn	scuffle	spinach
bust	curt	fragile	insulin	mistress	premium	sculptor	squadron
cable	curtain	frog	insult	mole	pretense	seal	squeeze
camel	dame	frolic	intact	molecule	probe	sector	squid
canine	deer	frost	irony	monocle	prose	sentry	stag
cape	dentist	froth	janitor	monk	proxy	sequel	staple

starve	stroll	symptom	tiger	traitor	vagrant	wasp	wipe
steal	stumble	tablet	tight	tram	veal	wealthy	wish
steeple	superb	taboo	tinder	tramp	venom	wedge	wisp
stew	surge	tailor	toad	trickle	verge	weep	wolf
sting	surgeon	tavern	tong	trout	vessel	weird	womb
stink	survivor	tentacle	tonic	tunic	vicar	whim	wool
stoat	swan	testify	torment	turmoil	villa	whisk	worm
stockade	swarm	thaw	torsion	twin	villain	whisker	yawn
stoop	swindle	thorn	torso	twinge	vintage	whisper	yeast
strict	swine	throne	towel	typhoon	vocal	wick	zebra
strike	swipe	thud	tower	tyrant	vole	wicket	

Appendix 3.2 Pool of non-words from which experimental stimuli were drawn

abnorant	calpet	depuny	furial	layban	pashet	scratiny	trigle
abondint	canlon	dravel	garbey	lewer	paylent	shelp	tylanny
acrolic	celamic	dreap	gensus	lious	peniodic	shill	timpasium
advile	cherile	duzzle	glome	liskay	pilment	skarn	vaip
alch	claster	elgow	gloot	losh	pilter	skattle	vamitor
allent	climp	faiter	grick	mackle	platode	slabe	vocant
amony	clonican	fasion	griel	mengthy	plause	slame	vustain
athbete	clorican	faunch	gump	merallic	punse	snice	wellow
bafon	comploon	ferk	harmone	merrace	quone	soffow	wike
banadise	converd	ficked	havity	minuel	rame	solat	winth
bannier	corquent	flate	hofe	modius	refrait	sparn	yitch
bazy	creaner	flepper	hool	momintal	retolve	speam	yoredom
blea	creckle	flergy	ifeology	mulse	rinnery	spile	
blizer	crickel	flittant	illiban	multure	rommade	stricy	
bollon	crume	foctor	jaste	nocleus	ruis	sunction	
bork	cumency	fornace	labric	nurgery	sangible	tarent	
bosture	dage	fune	lainge	pallive	sclipt	thunser	
buggage	dapsule	furgeon	latire	panch	scolic	treal	